Analyzing the molecular basis of plant root responses to the environment

Liam Walker

Doctor of Philosophy

University of Warwick
School of Life Sciences
September 2019

Table of contents:

List of abbreviations:

ABA – ABscisic Acid

AON – Autoregulation Of Nodulation

BP – Base Pairs

CCaMK – Calcium-CalModulin dependent protein Kinase

CHX – CycloHeXimide

CLE – CLAVATA/Embryo-surrounding region

DAI – Days After Inoculation

DAMP – Damage-Associated Molecular Pattern

DE – Differential Expression

DEX – DexaMethasone

DNA – Deoxyribose Nucleic Acid

EPS – Extracellular Polymeric Substances

ETI – Effector-Triggered Immunity

FACS – Fluorescence Activated Cell Sorting

GFP – Green Fluorescent Protein

GMO – Genetically Modified Organism

GO – Gene Ontology

GR – Glucocorticoid Receptor

GUS – β-GlUcuronidaSe

INTACT – Isolation of Nuclei TAgged in specific Cell Types

IT – Infection Thread

JA – Jasmonic Acid

PEP – Plant Elicitor Peptide

PRR – Pattern Recognition rReceptor

LCO - LipoChitoOligosaccharides

LPS – LipoPolySaccharide

LR – Lateral Root

LRR – Leucine-Rich Repeat

MAMP – Microbe-Associated Molecular Pattern

MAPK – Mitogen-activated protein kinase

miRNA – microRNA

MTI – MAMP-Triggered Immunity

NCR – Nodule Cysteine-Rich

NFR – Nod Factor Receptor

NIN – Nodule INception

NLR – Nucleotide-binding and Leucine-rich Repeats

NORK – NOdulation Receptor Kinase

NRE – Nitrate-Responsive Element

PNR – Primary Nitrate Response

PRR – Pattern Recognition Receptor

RFP – Red Fluorescent Protein

RNA – Ribose Nucleic Acid

ROS – Reactive Oxygen Species

RT-qPCR – Real Time quantitative Polymerase Chain Reaction

SA – Salicylic Acid

SCR – SCARECROW

SHR – SHORT-ROOT

SYMRK – SYMbiosis Receptor Kinase

TARGET – Transient Assay Reporting Genome-wide Effects of Transcription factors

TF – Transcription Factor

TML – Too Much Love

TRAP – Translating Ribosome Affinity Purification

WAK – Wall-associated kinase

List of figures and tables:

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Acknowledgements:

Although the work set out in this thesis is predominantly my own, I could not have achieved this without the help and advice of countless others. I am indebted to Gary Grant from the Phytobiology Facility for his work in providing me with materials to conduct my experiments and support for the growth and propagation of my plant material during this project. I would also like to thank Jumoke (Esther) Ogunniyi and Pélagie Ratchinski who assisted me with some of my labwork.

I would also like to thank all of the people who have helped me with the bioinformatic aspects of my project. Dr. Charlotte Rich provided me with invaluable guidance on identifying homologs of genes in other species and Luke Richards, Dr. Mark Walsh and Javier Antunez-Sanchez for their guidance and patience in dealing with my questions during the analysis of my RNAseq data.

I am particularly grateful to Dr. Ruth Schäfer and all the other members (especially Dr. Alonso Javier Pardal-Bermejo, Cantuğ Bar and Dr. Rob Maple) of lab C030 through the years who have been at hand to offer me advice of technical procedures during my time in the lab. I would also like to record my thanks to my second supervisor, Dr. Sascha Ott, and the members of my advisory panel, Dr. Steve Jackson and Dr. Jose Gutierrez-Marcos for their feedback throughout my project.

Finally, I wish to dedicate my final acknowledgements to the two people who have helped me the most during my studies. I wish to thank my supervisor, Dr. Miriam Gifford for giving me the chance to take on this project, her constant guidance and feedback on my experiments and her concern not just on my research but my wellbeing and career development. I also send my gratitude and admiration to my esteemed colleague Dr. Beatriz Lagunas for her patience and generosity with her time, her interest in my work and for her invaluable guidance over the years.

Declaration:

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

The results contained in Chapter 3: “Regulation of gene expression by nitrogen and rhizobia underpinning lateral root emergence in *Arabidopsis thaliana*” are part of a published work and as such contains contributions from others as detailed in the Author Contributions of the publication (Walker *et al*., 2017). Laboratory work for these experiments were conducted by Jesper Grønlund, Sanjeev Kumar, Jo Hulsman and Dhaval Patel. Ying Wang, Dafyd Jenkins and Clare Boddington were responsible for network inference, clustering and producing lists of differentially expressed genes from these experiments. Subsequent interpretation and analysis of this data is my own work.

Summary:

As sessile organisms, plants must be able to adapt to and exploit their environment in order to survive. A key aspect of this is the ability of plants to remodel their root system architecture in order to carry out the essential functions of providing anchorage and nutrient and water uptake from the surrounding soil. Soil typically contains a huge variety of microorganisms which will likely include species which are potentially harmful or beneficial to the plant, as well as a range of abiotic conditions.

One way in which plants can adapt their root systems in response to their environment is via the formation of new lateral roots. Lateral roots generally emerge perpendicularly to the primary root or other lateral roots and increase the surface area of the root system and the range of exploration. Genes involved in the regulation of lateral root formation in *Arabidopsis thaliana* were investigated by using fluorescence activated cell sorting over a timecourse. Gene expression changes over time in response to nitrogen application or *Sinorhizobium meliloti* inoculation, both of which are associated with the regulation of lateral root development, were investigated in cells of the pericycle, from which lateral roots derive, and an overlaying cell type, the cortex. Gene expression was found to be highly cell-type specific between the two cell types and this was conserved during environmental responses.

The formation of root nodules by legumes represents another quintessential example of a modification of the root to adapt to the environment. During conditions of nitrogen starvation, the plant can form structures on the root which can be colonized by symbiotic nitrogen-fixing bacteria in the soil called rhizobia. In the nodule, atmospheric nitrogen is reduced by the bacteria and utilized by the host plant. The intersection between plant defence responses and symbiosis was investigated in the model legume *Medicago truncatula* to try and identify genes involved in distinguishing rhizobia as symbionts rather than as pathogens. Putative novel markers of defence and symbiosis were identified that may underpin this transition.

Chapter 1: Introduction

Plant root architecture and the environment:

Plant roots are important for providing anchorage to the soil and facilitating nutrient and water uptake. In order to optimize their root systems for these functions, plants must be able to fine tune their root system architecture in response to their environment (Lynch, 1995). This need to fine tune is further complicated by the need to respond to a wide range of short-term biotic and abiotic stresses which affect resource allocation as well as growth. Plant root systems are capable of an exquisite level of plasticity and this serves as a mechanism by which they can adapt to their environment. As part of this mechanism, the plant must continue to integrate a multitude of diverse and often antagonistic signals when making developmental decisions. The molecular basis of this mechanism is the differential expression (DE) of genes; different environmental conditions will lead to the activation or repression of specific transcription factors (TFs) which in turn leads to different patterns of gene expression, the activity of which facilitates different developmental outcomes.

Much research has been carried out towards understanding the molecular basis of how plants regulate various developmental processes during environmental responses but this research is made difficult by the nature of these responses. Many of these responses occur over relatively long periods and involve different cell types maintaining vastly different transcriptional programs whilst still working in a concerted fashion.

Lateral root formation enables plants to adapt to their environment:

As sessile organisms, plants must source all the nutrients they need for growth from their local environment. The majority of nutrients plants require for growth are obtained from the soil so it is one of the primary functions of the root system to maximize the uptake of these. Plants meet this goal by ensuring their roots grow in the direction of nutrients. Growing from only a single (primary) root would reduce the likelihood of finding nutrients in addition to providing

extremely poor anchorage. Therefore, plants are able to induce branching in their root systems to maximize soil exploration and nutrient uptake. The formation of these branches, or lateral roots (LR), is a highly regulated process controlled by a diverse range of stimuli including nutrient status, abiotic stresses and even interactions with microorganisms in the rhizosphere that can both positively and negatively influence the morphogenesis of new LRs (Benkova and Bielach, 2010).

Stages of lateral root morphogenesis: LR formation involves forming an entirely new organ from an initially very small number of cells. LR initiation originates deep inside the root and thus requires co-ordinated action between multiple different root cell types for the root to eventually emerge through overlaying tissues. Because LR formation is best understood within the plant *Arabidopsis thaliana* (Figure 1.1), the following information is based on LR formation in this model. However, the mechanism appears to be highly conserved across other higher plant species such as *Medicago truncatula* (Herrbach *et al*., 2014).

LR formation can be considered to begin with the specification of a founder cell. Accumulation of the plant hormone auxin in a single cell of the xylem pole pericycle is sufficient to drive differentiation into a founder cell by activation of the transcription factor (TF) *GATA23*. (De Rybel *et al*., 2010; Dubrovsky *et al*., 2008). Cells which have adopted founder identity (as controlled by *GATA23* expression) then start to undergo changes. The nucleus migrates to a pole of the cell which then begins to divide asymmetrically, producing unevenly sized daughter cells and giving rise to a structure called the lateral root primordia (De Smet *et al*., 2006). At this stage, the process may be aborted by the alteration of cytokinin signalling which is able to inhibit founder cell division (Laplaze *et al*., 2007).

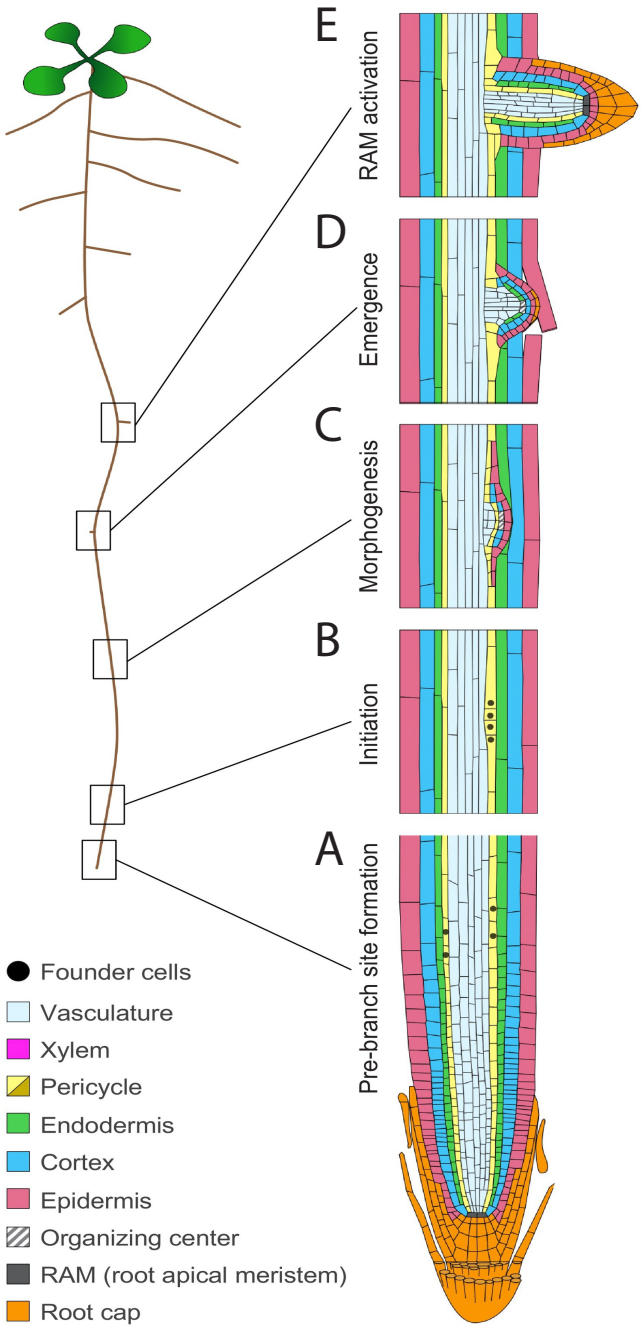

Figure 1.1: Stages of lateral root emergence in *A. thaliana*. Lateral root formation begins with the specification of founder cells (A) in the pericycle. These founder cells begin to divide (B) both radially and periclinally to form a lateral root primordia (C). Continued cell division lead to the expansion of the primordia and its emergence through overlaying tissues (D) before the lateral root primordia develops its own active meristem (E) and commences outgrowth into the surrounding environment. Adapted from Banda *et al* (Banda *et al*., 2019).

Broadly, the next stage of LR development is outgrowth. After several rounds of cell division, a dome shaped structure is formed as additional pericycle cells distort the structure of the overlaying tissues. Once the primordia is several cell layers deep it develops its own meristem (Laskowski *et al*., 1995). At this stage the emergence of a new LR is not inevitable as the LR primordia is still subject to regulation by the overlaying cells and stress signalling. The plant hormone abscisic acid (ABA) is responsible for inducing quiescence of LRs (De Smet *et al*., 2003). This may be mediated by the *MYB93* TF which is induced by ABA and expressed specifically in the endodermis overlaying LR primordia and serves to inhibit LR emergence (Gibbs and Coates, 2014; Gibbs *et al*., 2014).

Following outgrowth is emergence when the LR primordia irreversibly commits to becoming a LR and breaks through the overlaying tissues. Crossing the epidermis represents a particular challenge due to the presence of a lignin diffusion barrier called the Casparian strip that exists between endodermal cells. Selective breakages in the Casparian strip and shrinkage of endodermal cells allows the primordia to traverse the endodermis without affecting the integrity of overlaying cells and still protecting the vasculature from diffusion (Vermeer *et al*., 2014). Auxin signalling via the cortical and epidermal auxin transporter LAX3 (Auxin1-like protein 3) leads to loss of cross-linking between cells allowing the LR to grow through the now separated cells in these layers (Peret *et al*., 2009).

Now free from the mechanical constraints of the overlaying tissue, the new LR is able to grow outwards into the environment, enabling the plant to forage for resources. Although LRs predominantly arise from the primary root, it is possible for LRs to initiate from an existing LR, giving rise to second or even third order LRs. It should be noted that it is also possible, although rare, for roots functionally analogous to LRs to emerge from the shoot in some species, including *A. thaliana* (Gutierrez *et al*., 2012). These roots, called adventitious roots, are usually formed in response to mechanical or environmental stress (Bellini *et al*., 2014).

Factors influencing root system architecture: Plants integrate a diverse range of cues when regulating when and where to form new LRs. One of the

most potent external stimuli that influences root architecture is the form and availability of nutrients, particularly nitrogen in the form of nitrate. Proliferating roots towards areas rich in a key nutrient (foraging) is crucial for plants to adapt to their environment. The formation of new LRs is understood to be induced following a transition from deplete to replete nitrate (Linkohr *et al*., 2002; Walker *et al*., 2017; Zhang and Forde, 1998). A MADS box TF, *ANR1* (Arabidopsis Nitrate-Regulated 1), has been shown to be essential for orchestrating this response by stimulating lateral root emergence (Gan *et al*., 2005; Zhang and Forde, 1998). Conversely, LR initiation is repressed under low nitrate. As this phenotype is lost in a mutant of nitrate transporter *nrt2.1*, it seems NRT2.1 (Nitrate Transporter 2.1) mediates this inhibition, suggesting it has an additional function as a nitrate sensor (Little *et al*., 2005).

Water foraging and osmotic stress also influence root architecture significantly. Osmotic stress has been shown to repress (via the action of ABA) LR formation in *A. thaliana* (Deak and Malamy, 2005) and the cereals *Zea mays* and *Hordeum vulgare* (Babe *et al*., 2012). More recent work in the latter two species has suggested that the repression may be specific to LRs not in contact with water. This is via a response termed xerobranching which occurs when ABA accumulates in root sections near air spaces in the soil and prevents LR primordia initiation (Orman-Ligeza *et al*., 2018). LRs are known to preferentially emerge from the root towards water sources (hydropatterning). Hydropatterning is auxin-dependent but is unaffected in ABA mutants which suggests that it is a distinct process from the aforementioned water stress response (Bao *et al*., 2014). Subsequent work has shown this depends on *LBD16* (Lob domain-containing protein 16) expression in founder cells which itself depends on the TF *ARF7* (Auxin Response Factor 7); when roots are exposed to air SUMOlyation of ARF7 disrupts it binding to LBD16 and auxin signalling is disrupted (Orosa-Puente *et al*., 2018).

The effects of salinity on LR development depend on the severity of the salt stress. Mild salt stress seems to promote LR morphogenesis whilst more severe salt stress is inhibitory (Galvan-Ampudia and Testerink, 2011). The mechanism relating to salt perception is not well understood but LR initiation is mediated by auxin accumulation in the primordia as typical in unstressed

conditions (Zhao *et al*., 2011; Zolla *et al*., 2010) whilst higher salt concentrations induce accumulation of ABA and thus push the balance in favour of quiescence (Duan *et al*., 2013).

In addition to abiotic factors, biotic factors may also influence root plasticity. A retardation of systemic growth is a commonly observed phenotype of plants undergoing defence responses, possibly due to the plant reallocating resources from growth to defence. This is distinct from pathogenesis (for example a biotrophic pathogen diverting resources from the plant) as the same phenotype is observed in response to treatment with the immunogenic peptide flg22 which activates innate immunity but is harmless to the plant (Gomez-Gomez *et al*., 1999). Phenotypes have also been reported in some non-pathogenic plant-microbe interactions; for instance, treatment with the rhizobial symbiont *Sinorhizobium meliloti* triggers the emergence of new LRs at the expense of elongation in *A. thaliana* (Walker *et al*., 2017).

Nodulation is a symbiosis between legumes and nitrogen-fixing bacteria:

The formation of root nodules by leguminous plants represents another example of plants altering their root systems in response to their environment. The family Fabaceae (legumes) is set apart for almost all of the rest of the plant kingdom by the ability of most of its family members to enter symbiosis with soil-dwelling nitrogen fixing bacteria called rhizobia. During this symbiosis (nodulation) the bacteria physically colonize specialized structures on the plant root called nodules where they reduce atmospheric nitrogen into forms that can be utilized by the host plant (Oldroyd, 2013).

Nodule organogenesis requires co-ordination between the epidermis and cortex: Similarly to lateral root formation, nodule morphogenesis is a process that requires the concerted action of multiple cell types (Figure 1.2). Nodulation typically exhibits a high degree of species specificity with different rhizobial strains showing preference to colonizing particular host species and vice-versa (Albrecht *et al*., 1999). Although the majority of these interactions are characterized by a similar set of events, nodulation is best characterized

in the model legumes *Medicago truncatula* and *Lotus japonicus* so subsequent information is predominantly based on the events from these models.

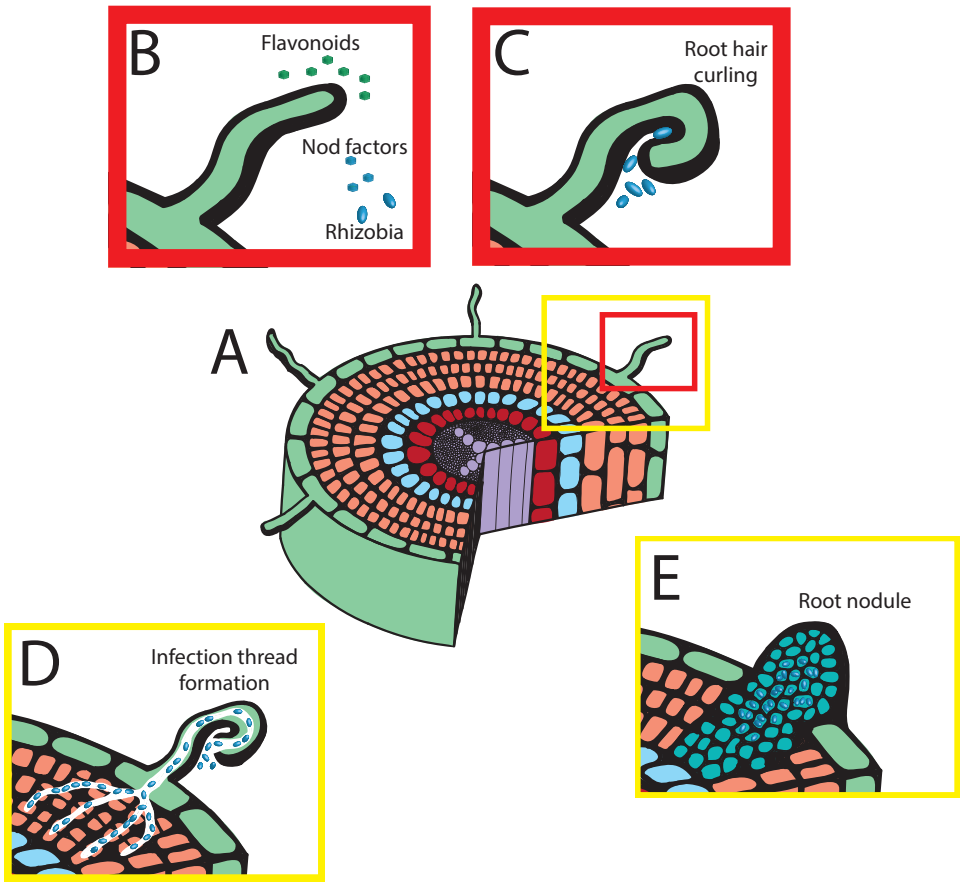

Figure 1.2: Stages of indeterminant nodule formation. (A) Cross-section of a plant root showing individual cell types comprising different layers of the root; stele (purple), pericycle (red), endodermis (blue), cortex (orange) and endodermis (green). (B) Nodulation is initiated when, in nitrate-deplete conditions, the plant root secretes flavonoids into the rhizosphere. These are detected by nearby rhizobia which respond by secreting Nod factors and are chemoattracted into close proximity of the plant root hairs of epidermal cells. (C) Nod factor perception by the plant results in calcium ion spiking in epidermal cells and the induction of a suite of nodulation-specific signaling and transcription factor genes that lead to both root hair curling that encircles the rhizobia, and initiation of cell division within underlying cortical cells. (D) Trapped bacteria are able to invade the root through a structure called an

infection thread leading to invasion of underlying cortical cells. Rhizobial invasion triggers rapid cell division in the cortex. As the developing nodule continues to grow it breaks through the overlaying endodermis. (E) Rhizobia differentiate into specialized nitrogen-fixing bacteroids and assimilate atmospheric nitrogen for use by their host plant, receiving carbon-based compounds derived from photosynthesis in return. The nodule continues to grow and can eventually exceed the diameter of the root it derived from.

Nodulation begins with signalling cross-talk between legume root hair cells and rhizobia in the soil secreting flavonoids and Nod factors respectively (described in more detail below). This signalling culminates in chemoattraction of the bacteria into close proximity of epidermal root hairs (Bauer and Caetano-Anollés, 1990). As a result of Nod factor-induced localized inhibition of growth at root hair tips, root hairs are induced to curl, creating a pocket in which rhizobia may become trapped (Esseling *et al*., 2003). Rhizobia then gain entry into the root hair by localized cell wall degradation which is thought to be host-mediated (Xie *et al*., 2012). Cytoskeletal rearrangements occur in the root hair, giving rise to a structure called an infection thread (Timmers *et al*., 1999). By dividing towards the root, the bacteria in the IT are able to traverse the infection thread which ultimately extends into underlying cortical cells (Gage, 2004).

Concurrent with root hair curling and infection thread formation in the epidermis is de-differentiation and division of the first two layers of underlying cortical cells giving rise to a nodule primordia which will become the nodule meristem (Patriarca *et al*., 2004). The nodule continues to grow in size and can eventually host as many as 10^9 bacteria (Downie, 2014). However, there is a distinction between legumes in which this meristem is transient (determinant nodulators) or maintained (indeterminant nodulators). The nodules of indeterminately nodulating species, such as *M*. *truncatula*, are therefore able to persist indefinitely whilst determinant nodules formed by species including *L*. *japonicus* eventually senesce (Gibson *et al*., 2008).

Bacteria in the nodule primordia are enclosed within a membrane by endocytosis, becoming an organelle called the symbiosome (Brewin, 2004). Bacteria in the symbiosome usually differentiate into specialized nitrogen-fixing bacteroids, losing their ability to replicate in the process (Oke and Long,

1999). Inside the nodule, bacteria use the enzyme nitrogenase to break down atmospheric dinitrogen into ammonia which is provided to the host plant whilst anaerobic conditions in the symbiosome protect the enzyme from oxygen (Hunt and Layzell, 1993). Accumulated ammonia is used by the host plant for nitrogen nutrition whilst the bacteria obtain fixed carbon from photosynthesis in exchange (Kaschuk *et al*., 2009). The level of ammonia/carbon exchange is highly regulated over the course of nodule development.

Molecular activation of the nodulation pathway: During times of nitrate paucity, compounds called flavonoids are exudated from the root of the host plant into the rhizosphere. These flavonoids are able to diffuse across rhizobial membranes and into the cell (Fisher and Long, 1992). Upon flavonoid perception, rhizobia activate transcription of a set of genes (Nod genes) orchestrated by nodulation protein D (*nodD*) binding to the *nod* box of these genes. The principle result of this transcription programme is the production and secretion of lipochitooligosaccharides (LCO) compounds called Nod factors (Schlaman *et al*., 1992). Although the primary function of flavonoids during the symbiosis is the activation of Nod factor production in rhizobia, they are also essential to endogenous processes in legumes including repression of auxin transport and activation of genes involved in the nodulation pathway (Subramanian *et al*., 2007).

Nod factors are crucial in reprogramming the host plant to begin forming nodulation (Figure 1.3). Nod factors are recognized by RLK proteins homologous to the *A. thaliana* chitin receptor *CERK1* (Chitin Elicitor Receptor Kinase 1). The Nod factor receptors (NFR) *NFR1* and *NFR5* were first identified in *L. japonicus* (Madsen *et al*., 2003; Radutoiu *et al*., 2003) and their orthologs, *LYK3* (Smit *et al*., 2007) and *NFP* (Nod factor perception) (Amor *et al*., 2003) respectively were later found in the *M. truncatula* genome. Unlike most RLKs, NFR5/NFP (Fliegmann and Bono, 2015; Mbengue *et al*., 2010) lacks an active kinase domain, suggesting that both receptors must interact for Nod factor perception with NFR5/NFP functioning as a signalling receptor and NFR1/LYK3 as an entry receptor. The exact mechanism of by which these receptors function is not fully understood but they are known to co-localize at the plasma membrane (Moling *et al*., 2014). Some targets immediately

downstream of these receptors are beginning to be identified. LYK3 is able to phosphorylate LYR3 (Fliegmann and Bono, 2015) although the significance of this interaction is unclear. NFR5 has been shown by co-immunoprecipitation to associate with the cytoplasmic kinase NiCK4 (NFR5-interacting Cytoplasmic Kinase) which is then able to phosphorylate NFR1 and NFR5 (Wong *et al*., 2019). Nod factor signalling must be maintained through subsequent infection thread formation and also maintained in nodules for the symbiosis to be successful as later stages of nodulation are also induced by Nod factor impression (Timmers *et al*., 1998).

Although Nod factors are usually considered indispensable for nodulation there are examples which contradict this. For instance, some species of legume are nodulated in the absence of Nod factor (and thus without an infection thread) when the bacteria directly invade through naturally-occurring gaps in the epidermis via a process termed 'crack entry' (Sprent, 2008). However, examples of invasion without infection threads that still depend on Nod factors also exist (Acosta-Jurado *et al*., 2016; Liang *et al*., 2019). Some rhizobia that normally do depend on Nod factors for infection are still able to induce nodulation in their absence; a Nod factor-deficient mutant of *Bradyrhizobium elkanii* was still able to, possibly also via crack entry, weakly nodulate soybean and this interaction was dependent on the intact expression of a type III secretion system by the bacteria (Okazaki *et al*., 2013). More rarely, some rhizobia are able to directly infect epidermal cells via a mechanism termed intracellular infection. This process is distinct from crack entry as it occurs independently of natural gaps, with the bacteria using other methods to gain access, possibly by degrading epidermal cell walls (Ibanez *et al*., 2017). Finally, some mutants are able to be nodulated in the absence of Nod factors or even rhizobia. For example, constitutive activation of the CCaMK protein (Calcium/calmodulin-dependent kinase, discussed below) by removal of its autoinhibitory domain is sufficient to induce spontaneous nodule organogenesis, giving rise to non-functional nodules devoid of rhizobia (Gleason *et al*., 2006; Tirichine *et al*., 2006).

Downstream of NFRs, are more proteins involved in the nodulation pathway. The first of these is the protein *SYMRK* (Symbiosis receptor kinase, *L. japonicus* (Stracke *et al*., 2002))/*NORK* (Nodulation receptor kinase, *M.*

truncatula; also called Doesn't Make Infections 2 (DMI2) (Endre *et al*., 2002)). It operates downstream of Nod factor receptors and is responsible for transducing Nod factor perception into nuclear calcium oscillations (Charpentier and Oldroyd, 2013). This Ca^{2+} spiking is dependent on the ion channels *CASTOR* and *POLLUX* (Charpentier *et al*., 2008) and the nucleoporin *NUP133* (Kanamori *et al*., 2006) in *L. japonicus* or DMI1 in *M. truncatula* (Peiter *et al*., 2007).

The calcium ion flux is detected by the nuclear-localized *CCaMK* in *L. japonicus* or the orthologous *DMI3* in *M. truncatula* (Weidmann *et al*., 2004). Upon binding calcium, the resulting conformational change exposes a threonine residue which is then phosphorylated, activating the protein (Singh *et al*., 2012). CCaMK/DMI3 is found in complex with its target CYCLOPS/IPD3 (Messinese *et al*., 2007; Yano *et al*., 2008) which it then phosphorylates.

The active CCaMK/DMI3 and CYCLOPS/IPD*3* complex activates downstream targets that in turn induce nodulation-related gene transcription. The GRAS family TFs *NSP1* (Smit *et al*., 2005) and *NSP2* (Kalo *et al*., 2005) are the best characterized of these nodulation regulatory genes. The two proteins dimerize and activate expression of both early nodulin genes and additional TFs (Hirsch *et al*., 2009). One such component is the essential *ERN1*/*ERN2* complex (Cerri *et al*., 2012; Hirsch *et al*., 2009; Middleton *et al*., 2007) which is implicated in infection thread formation (Kawaharada *et al*., 2017). Nodule inception (*NIN*) is also a key regulator of nodulation (Marsh *et al*., 2007) and has recently been found to be highly conserved not just in legume species, but also non-legumes that are able to nodulate including *Parasponia andersonii* and *Casuarina glauca* (Clavijo *et al*., 2015; van Velzen *et al*., 2018). *NIN* is known to have multiple roles during nodulation, including activation of the *CRE1* cytokinin receptor. Cytokinin signalling is known to promote nodule inception (Gamas *et al*., 2017) in the cortex but actually antagonizes processes associated with further infection in the epidermis, suggesting it has a role in feedback inhibition of the nodulation pathway (Vernie *et al*., 2015; Yoro *et al*., 2014).

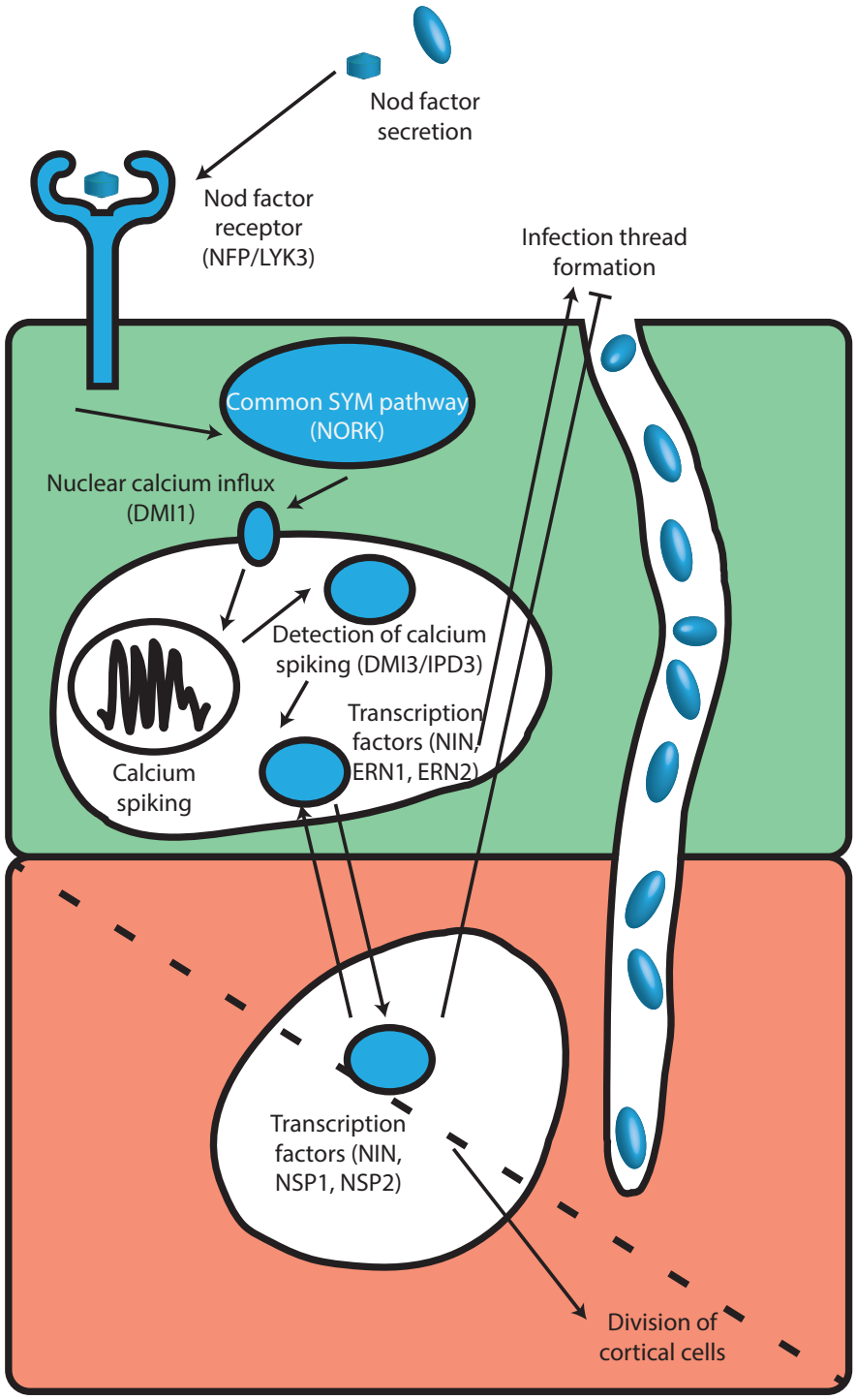

Figure 1.3: Key genes involved in activation of the nodulation pathway in *M. truncatula*. Initial perception of rhizobial Nod factors is mediated by the NFP/LYK3 receptor complex in the epidermis. The receptor complex activates NORK leading to opening of the DMI1 ion channel and nuclear calcium influx. Calcium influx activates the DMI3/IPD3 complex which leads to the activation

of transcription factors associated with nodulation. Transcription factor activity in the epidermis enables infection by rhizobia whilst driving cell division in the cortex and later inhibiting further infection thread formation. Adapted from Lagunas *et al* (Lagunas *et al*., 2015).

Regulation of the nodulation pathway: Nodulation incurs significant costs for legumes in terms of resources and is accordingly very tightly regulated to maximize the benefits to the host. Similarly, rhizobial activity is controlled according to resource allocation to the symbiosomes. Control of nodulation does not simply consist of a binary passage or blockage of nodulation; legumes are able to finetune nodule numbers to best balance their energy needs. Nodulation is able to be moderated at all stages of the symbiosis including once active nodules have developed (Ferguson *et al*., 2019).

The main benefit legumes derive from the symbiosis is available nitrogen, thus it does not make sense for the plant to form nodules during a surfeit of nitrogen. Therefore, nodulation is primarily regulated through the availability of nitrogen although this shares mechanisms with the autoregulation of nodulation pathway (Okamoto and Kawaguchi, 2015). It has long been known that nodule formation is inhibited in high nitrate concentrations but the mechanisms behind this are not well understood and may vary between legume species (Mortier *et al*., 2012; Omrane and Chiurazzi, 2009). Plants grown in a high nitrate environment remain resistant to rhizobial colonization for several days after transfer to low nitrogen (Omrane *et al*., 2009). Flavonoids may be a possible mechanism controlling this; the expression and diversity of flavonoids can be reduced in replete nitrate conditions which could consequently reduce plant-rhizobia interactions (Cho and Harper, 1991). Both nitrate and ammonia act to inhibit rhizobial perception at some point between root hair curling and cortical cell division (Barbulova *et al*., 2007; van Noorden *et al*., 2016). Additionally, the size and activity of existing nodules is also rapidly reduced in higher nitrate conditions and this is likely mediated by cutting off the supply of photosynthates to the nodule (Fujikake *et al*., 2003).

In addition to direct nitrate-mediated inhibition, legumes also use a feedback inhibition pathway called autoregulation of nodulation (AON). AON

involves induction of CLAVATA3/embryo-surrounding region (CLE) peptides by rhizobia which inhibit additional infection (Reid *et al*., 2011). In *M. truncatula* these peptides are called CLE12 and CLE13 (Mortier *et al*., 2010) with close orthologs identified in other legume species including *L. japonicus* (Nishida *et al*., 2016; Okamoto *et al*., 2009), *Glycine max* and *Phaseolus vulgaris* (Ferguson *et al*., 2014). In *L. japonicus*, the production of CLE peptides is known to occur via the previously discussed TF *NIN* (Soyano *et al*., 2014). Functional CLE peptides are first synthesized as much longer pre-cursors before they are cleaved to peptides 12-13 amino acids in length (Hastwell *et al*., 2015) and then further chemically modified via arabinosylation (Corcilius *et al*., 2017; Hastwell *et al*., 2019; Okamoto *et al*., 2013). The mature CLE peptides are then transported by the vasculature to the shoots where the CLE receptor is expressed (Magori and Kawaguchi, 2009). In *M. truncatula* the CLE receptor is called *SUNN* (Super numerary nodules) and its mutation results in a hypernodulation phenotype as the plant loses its ability to regulate nodule number (Schnabel *et al*., 2005). Analogous receptors have been identified in *Pisum sativum*, *L. japonicus* (Krusell *et al*., 2002; Nishimura *et al*., 2002), *Glycine max* (Searle *et al*., 2003) and *P. vulgaris* (Ferguson *et al*., 2014). Upon activation of the CLE receptor an unknown signal originates in the shoots and travels to the roots, again via the vasculature, where it is able to antagonize further nodule development. Very little is known about the subsequent steps aside of studies in the *L. japonicus* hypernodulating mutant *too much love* (tml). TML is known to act specifically in the root downstream of CLE receptor *HAR1* (Hypernodulation Aberrant Root formation 1) and likely functions by targeting unknown protein(s) for degradation by the 26s proteasome pathway (Magori *et al*., 2009; Takahara *et al*., 2013).

The plant also responds to and integrates hormonal cues in order to inhibit nodulation when conditions are not conducive to nodulation. The plant hormone ethylene is a potent inhibitor of nodulation likely due to its role in signalling for other stresses (Stougaard, 2000). The defence-associated hormone jasmonic acid (JA) has been demonstrated to inhibit nodulation in *M. truncatula* (Sun *et al*., 2006) and *L. japonicus* (Nakagawa and Kawaguchi, 2006) seemingly by inhibiting the early nodulation apparatus but acting independently of any direct effect on rhizobial wellbeing. The role of another

defence-associated hormone, salicylic acid (SA), is less clear. Although traditionally thought to be an inhibitor of nodulation only in indeterminately nodulating species (van Spronsen *et al*., 2003), external SA application has been shown to inhibit nodulation in the determinant nodulator *L. japonicus*. Reducing SA accumulation in these plants by expression of the bacterial salicylate hydroxylase gene *nahG* also leads to increased nodulation (Stacey *et al*., 2006).

Finally, various environmental stresses have negative effects on the fitness of both symbiotic partners and these lead to inhibition of nodulation. This allows the plant to conserve resources for coping with the stress although the mechanisms underpinning this are even less well understood than for other forms of inhibition. Low soil pH leads to reduced nutrient availability and increased accumulation of toxic metal ions which reduce the effectiveness of nitrogen fixing symbiosis. In *Glycine max*, systemic repression of nodulation acting via the CLE receptor from the AON pathway has been reported thus providing a mechanism for controlling nodulation in acidic soils (Lin *et al*., 2012). High soil salinity severely harms rhizobial nitrogen fixing efficiency and many legumes exhibit reduced nodule numbers and nitrogen fixation as a coping mechanism (Manchanda and Garg, 2008). Drought stress also leads to reduced nodulation even when the plant is able to maintain photosynthate supply to nodules, suggesting this may simply reflect reduced demand for nitrates due to reduced growth and that drought stress is able to feed into the AON pathway (Streeter, 2003).

Determinants of host-symbiont specificity during nodulation: Nodulation represents a significant cost to legumes in terms of providing fixated carbon to bacteroids in nodules. An optimal strategy is for the plant to only participate in symbiosis with bacteria that fix nitrogen efficiently in return. However, a lifestyle closer to parasitism (cheating) may be more beneficial from the perspective of the bacteria if it able to sequester carbon whilst providing little or no nitrogen fixation in return. Cheating is especially a consideration in some legume-rhizobia interactions were bacteroids are terminally differentiated and therefore fixing nitrogen occurs a huge fitness cost to the rhizobial population as a whole (Denison and Kiers, 2004). Therefore, it is important for legumes

not just to be able distinguish between commensals, symbionts and pathogens but also to differentiate between individual symbionts on the basis of their compatibility with the plant (Clua *et al*., 2018).

The first opportunity to discriminate between symbiotic partners is during the initial signalling events between the plant and rhizobia. Across the legume family there is enormous diversity in flavonoids although not all are involved in nodulation (Cooper, 2007). Flavonoids serve as the primary determinant of rhizobial host range based on their ability to activate *NodD* expression in the bacteria (Liu and Murray, 2016; Wang *et al*., 2012). This has been demonstrated elegantly using a *NodD* null mutant of *S. meliloti* transformed with *NodD* plasmids from other rhizobia species. Different combinations of flavonoids were required to induce Nod gene expression depending on the source of *NodD* complementation (Peck *et al*., 2006).

The ability of the host plant to recognize Nod factors during the initial cross-talk between the two parties is also a determinant of host range as subsequent events require expression of nodulation-related genes in the plant. Although Nod factors share the same basic structures, they can be extensively modified by the bacteria which has given rise to diversity of Nod factors across rhizobial species (Long, 1996). Removing the *nod* genes of *Rhizobium leguminosarum* abolished its ability to form nodules with its natural host *Trifolium refens*. Subsequent transfer of a plasmid bearing the *nod* genes from *S. meliloti* permitted symbiosis with *Medicago alfafa*, a natural host of the donor rhizobia (Debelle *et al*., 1988). This applies equally to the host plant; transformation of *M. truncatula* with Nod factor receptors from *L. japonicus* facilitates symbiosis with rhizobia which are normally not compatible including *Mesorhizobium loti* (Radutoiu *et al*., 2007).

Plant defence responses, or rather the ability to overcome these, is also thought to play a role in rhizobial compatibility. Surface and exopolysaccharides of rhizobia are often required for suppression of plant defence responses. Mutant strains of *S. meliloti* that do not produce the exopolysaccharide succinoglycan are not able to initiate infection threads during symbiosis with *M. alfafa* (Cheng and Walker, 1998). Some rhizobia are known to use type III or IV secretion systems to deliver effector proteins into target cells and aid infection. It has been theorized that some legumes may

have evolved resistance (R) proteins that detect these effectors and activate defences, thus serving as another mechanism of control of host range (Deakin and Broughton, 2009). This is exemplified by the *Rj2* and *Rfg1* genes in *Glycine max* which encode R proteins that block nodulation with certain strains of *Bradyrhizobium japonicum* and *Sinorhizobium fredii* respectively (Yang *et al*., 2010).

Although host/symbiont selection typically occurs prior to infection, mechanisms exist to select for preferential rhizobia post-infection. For instance, some legumes species including *G. max*, *M. alfalfa* and *P. sativum* are able to 'punish' rhizobia in nodules that fail to fix nitrogen by cutting off the supply of photosynthates to that nodule or altering the oxygen permeability of the nodule (Kiers *et al*., 2003; Oono *et al*., 2011).

Nodule cysteine-rich (NCR) peptides are also thought to be important determinants of host range in some indeterminant nodule forming legumes. The *M. truncatula* genome encodes more than 700 NCR peptides which are short sequences of amino acids related to antimicrobial defensins (Pan and Wang, 2017). *S. meliloti* strain Rm41 is able to form nitrogen fixing nodules in the *M. truncatula* ecotype DZA315 but not A17. This has been shown to be at least partially due to each ecotype possessing different versions of one NCR, *NFS2*, which has antimicrobial activity in only the latter case (Wang *et al*., 2017). Interestingly, mutants deficient in another NCR family member, *NCR169*, fail to fix nitrogen in nodules infected with either *S*. *meliloti* 1021 or *S*. *medicae* 419 (Horvath *et al*., 2015). This suggests that despite the very large number of family members, a single peptide can still be functionally non-redundant and that they have functions beyond simply supressing the populations of less favoured rhizobial species. *S. meliloti* B800 forms nitrogen fixing nodules in *M*. *truncatula* ecotype A17 but deficient ones in ecotype A20. This *S. meliloti* strain possesses the peptidase *hrrP* that is able to cleave some NCR peptides. It is possible that the phenotype in the former case is due to A17 possessing NCRs not targeted by this peptidase allowing it to keep bacteroids differentiated at the expense of their proliferation (Price *et al*., 2015).

Plant possess multiple layers of defences to protect them from pathogens:

Plants can be attacked by pathogens that try to obtain nutrients from living (biotrophic pathogens) or dead (necrotrophic pathogens) cells. Plants rely entirely on innate immunity to protect themselves from these threats. This immune system is complex and consists of multiple layers that act to minimize harm to the plant at different stages of pathogenesis (Figure 1.4). Defence responses arising from immunity from the plant can be either local to sites of invasion and wounding or systemic. (Boller and Felix, 2009). The components and regulation of these pathways is best understood in *A. thaliana* so most of the information in this section is based on this model. However, it is thought to be highly conserved across higher plants (Staal and Dixelius, 2007).

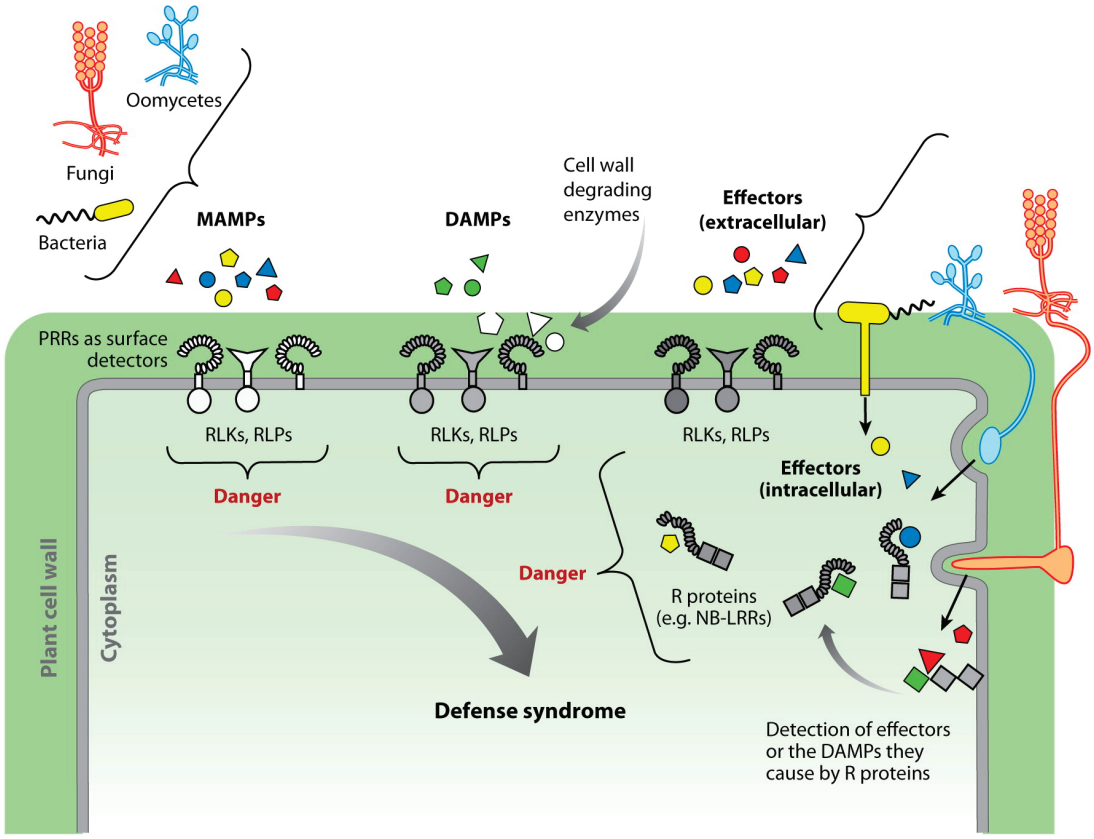

Figure 1.4: Mechanisms of plant defence. Receptors at the cell surface have evolved to recognize features of microbes (MAMPs) such as flagellin and

chitin and activate pathways leading to defence responses upon their perception. Pathogens have been found to secrete effectors into plant cells to enhance virulence of interfere with the immune response. These effectors, or symptoms of their activity, may be recognized by cytoplasmic receptors which also can activate defence responses. Another class of receptors exist that recognize endogenous signals associated with wounding of the plant (DAMPs) which are also able to trigger immunity. Adapted from Boller and Felix, 2009.

MAMP-triggered immunity recognizes chemical signatures of microorganisms that may be associated with pathogens: The first element of plant immunity is MAMP-triggered immunity (MTI). MTI consists of a range of pattern recognition receptors (PRRs) that identify conserved features of bacteria and other microbes, pathogenic or otherwise. Recognition of ligands by these PRRs induces an intracellular signalling cascade that results in protective immunity in most circumstances.

Bacterial flagellin has long been known to induce defence responses in *A. thaliana* and this is mediated by the PRR *FLS2* (Flagellin-sensitive 2) in *A. thaliana* (Gomez-Gomez and Boller, 2000). A specific 22 amino acid epitope (flg22) contained within flagellin acts as a portent MAMP and is responsible for binding FLS2 (Chinchilla *et al*., 2006) and that FLS2 has to form heterodimers with *BAK1* to elicit downstream signalling (Chinchilla *et al*., 2007). Chitin from fungal cell walls is also a MAMP and this is detected by the *CERK1-LYK5* PRR complex (Cao *et al*., 2014; Miya *et al*., 2007; Petutschnig *et al*., 2010). The highly conserved bacterial elongation factor EF-Tu also has MAMP activity and its recognition is conserved in Brassicaceae. It is perceived by the receptor *EFR* in *A. thaliana* (Zipfel *et al*., 2006).

DAMP-triggered immunity of plants responds to endogenous signals indicative of pathogen activity: In the event that MTI is not successful in restricting attack by a pathogen, another layer of immunity revolves around the detection of host-derived molecules called damage-associate molecular patterns (DAMPs). These elicitors are released during cell damage or death and therefore serve as an additional mechanism for activating defence responses (Yamaguchi and Huffaker, 2011).

The plant elicitor peptides (PEPs) are a family of DAMPs consisting of 7 peptides in *A. thaliana* and is thought to be conserved in many plant species (Huffaker *et al*., 2006). *A. thaliana* possesses two receptors for PEPs called *PEPR1* and *PEPR2* and both mutant studies and binding assays show partial redundancy between them. *At*PEP treatment activates defence responses and enhances disease resistance to pathogens. (Krol *et al*., 2010; Yamaguchi *et al*., 2010; Yamaguchi *et al*., 2006).

DAMP perception seems to activate immunity by the same pathways as MAMP perception. This suggests that DAMPs have evolved to increase the diversity of mechanisms by which plants can detect danger and to amplify existing immune responses, increasing the likelihood of a protective immune response (Yamaguchi and Huffaker, 2011). This theory is supported by induction of MTI leading to enhanced transcription of both PEP pre-cursors and PEP receptors (Ross *et al*., 2014; Yamaguchi *et al*., 2010).

Effector-triggered immunity counteracts attempts by pathogens to subvert the host immune system: Plant-pathogen interactions represent what is often termed an evolutionary arms race between the host and microbe. Some pathogens use proteins called effectors to circumvent host immunity and enhance virulence. Plants have in turn developed receptors termed resistance (R) proteins that can detect either the presence or activity of these pathogen effectors. This system is called effector-triggered immunity (ETI) and represents another key layer of plant innate immunity (Cui *et al*., 2015).

Pathogen effectors are diverse and so R proteins must be exquisitely specific, often recognizing the effector of a single species or even strain of pathogen. These R proteins belong to a family of receptors conserved throughout the plant kingdom called nucleotide-binding domain and leucine-rich repeats (NLR/LRR) proteins (Jacob *et al*., 2013). There are two modes of action by which NLRs may recognize effectors. Direct recognition occurs when the cognate ligand of an R protein is a bacterial effector and indirect recognition describes the perception of modification to endogenous proteins as a result of pathogen effector activity rather than the effector itself (Rafiqi *et al*., 2009). In either case, the receptors are thought to function as dimers although it is unclear if this is their native form or if this is induced by ligand

recognition. Active receptors then instigate transcriptional reprogramming, leading to a defence response (Cui *et al*., 2015). The complex co-evolution of host and pathogen is reflected in ETI; just as a plant might evolve an R protein recognizing an effector, the pathogen may then evolve an effector to supress the activity of that protein and this new effector could also become a target for the plant. In some cases, R proteins actually enhance the virulence of pathogens, especially in the case of necrotrophic pathogens which depend on host cell death to source nutrients (Lorang *et al*., 2007).

An example of direct recognition is provided by the oomycete *Hyaloperonospora arabidopsidis* effector *ATR1* and its corresponding NLR *RPP1*. The LRR domain of *RPP1* binds to the effector and activates an immune response (Krasileva *et al*., 2010). The *RIN4*/*RSP2*/*RPM1* model of indirect effector recognition provides an example of the potential complexity of ETI. The product of the *RIN4* gene is involved in signalling the closure of stomatal openings, a barrier to leaf pathogen entry (Liu *et al*., 2009a; Liu *et al*., 2009b). The effector proteins *AvrB* and *AvrRpm1* of the pathogen *Pseudomonas syringae* DC3000 phosphorylate and inactivate *RIN4* (Kim *et al*., 2005b; Mackey *et al*., 2002). *RPM1* confers resistance to these two effectors, likely through detecting these modifications and activating immune responses (Mackey *et al*., 2002). Another *Pseudomonas syringae* effector, *AvrRpt2*, is able to cleave and inactivate *RIN4* (Kim *et al*., 2005a; Mackey *et al*., 2003). Expression of the *RPS2* gene by the host plant confers resistance to *AvrRpt2* and is able to associate with *RIN4*. *RSP2* signalling is only induced in the absence of *RIN4* suggesting that it responds to the elimination of this protein rather than the effector itself (Axtell and Staskawicz, 2003).

Signal transduction from the perception of danger to the expression of defence-related genes: Plant immunity integrates a wide range of cues at the recognition stages but the responses downstream of these receptors are less specific. There are two main defence pathways in plants based around phytohormone signalling; the salicylic acid pathway (SA) is associated with defence against biotrophic pathogens and the jasmonic acid (JA)/ethylene pathway is usually active against necrotrophic pathogens. Both pathways are usually considered to be mutually antagonistic (Bari and Jones, 2009).

The activity of mitogen-activated protein kinases (MAPK) is crucial in transducing signals between active receptors and downstream TFs. MAPKs act at the bottom of cascades, usually depending on activating phosphorylation from a MAPK kinase (MAP2K) which may in turn require activation by an upstream kinase (MAP3K) (Pitzschke *et al*., 2009). The *A. thaliana* genome encodes 60 MAP3Ks, 10 MAP2Ks and 20 MAPKs (Rasmussen *et al*., 2012).

The best characterized MAPKs are *MAPK3*, *MAPK4* and *MAPK6*. *MAPK3*/*MAPK6* are thought to function together at the end of one cascade (Ren *et al*., 2008) and *MAPK4* in another (Gao *et al*., 2008). All of these are strongly linked to co-ordination of immunity in both PTI and ETI responses. For instance, all three are induced by flagellin (Bethke *et al*., 2009; Droillard *et al*., 2004) whilst MAPK3 is induced by purified effectors from the oomcyetes *Phytophthora parasitica*, *Pythium aphanidermatum* and *Phytophthora sojae* (Qutob *et al*., 2006), MAPK6 by the fungal pathogen *Alternaria brassicae* (Kannan *et al*., 2012) and MAPK4 and MAPK6 are induced by harpin (an immunogenic compound) secreted by *Psuedomonas syringae* (Desikan *et al*., 2001).

Activated MAPKs are then responsible for regulating TFs to co-ordinate the immune response. The best studied of these belong to the *WRKY* family of TFs. Following initial activation, some *WRKY* TFs are able to up-regulate their own expression and that of other family members by cross-regulation (Eulgem and Somssich, 2007) Both MAPK4 and MAPK3/MAPK6 cascades are known to activate WRKY33 (Andreasson *et al*., 2005; Mao *et al*., 2011). Interestingly, in the former case, the mechanism of WRKY33 activation is not phosphorylation, but release from complex with MAPK4/MKS1 when the MAPK is phosphorylated (Qiu *et al*., 2008). *wrky33* mutants do not form protective immune response to the necrotrophic fungal pathogen *Botrytis cinereal* and *WRKY33* overexpression results in increased resistance. Conversely, the response to *Pseudomonas syringae* is unaffected with over-expressing lines actually showing enhanced susceptibility (Zheng *et al*., 2006). This highlights the antagonism between the necrotrophic and biotrophic pathogen defence pathways.

Defence responses leading to protective immunity: Upon activation of TFs in the innate immune pathway, cells will begin transcriptional reprogramming towards immunity and this cumulates in defence mechanisms that lead to blockage of entry or denial of resources to pathogens.

One of the most immediate outcomes of defence activation is the production of reactive oxygen species (ROS). This results in an oxidative burst that prevents colonization by many pathogens. Although the high levels of ROS are themselves likely to be harmful to pathogens it is thought that localized cell death and cross-linking of host proteins attributable to oxidative burst is the primary mechanism of resistance (Bolwell, 1999).

Many plants are capable of the production of antimicrobial secondary metabolites called phytoalexins. In *A. thaliana* the most important phytoalexin is camalexin (Bednarek *et al*., 2011). Camalexin is toxic to some bacteria and fungi, likely by disrupting cell wall integrity (Glawischnig, 2007). Although such proteins have not yet been identified in *A. thaliana*, some crop species have been reported to secrete antimicrobial enzymes such as chitinase and β-1,3-glucanase which degrade fungal cell walls (Mauch *et al*., 1988).

Cell wall reinforcement is another defence mechanism effective against some pathogens. Deposition of the polysaccharide callose is one of the hallmarks of plant defence responses. Cell walls can be impregnated with callose, lignin and extensin proteins to protect healthy cells from invasion and to seal off infected cells (Hematy *et al*., 2009).

Fluorescence activated cell sorting facilitates the study of complex developmental processes in plants:

A key aspect of multicellularity is that it enables differentiation of cells into specialized types within organisms. These different cell types may then take on distinct functions allowing far greater complexity as the concerted action of many different cell types can give rise to tissues and organs which can in turn de-differentiate in specific locations to form new organs such as lateral roots and nodules. This specialization is underpinned by differential expression and activity of genes.

As discussed previously, responses to both abiotic and biotic (symbionts and pathogens) stimuli involve a high degree of cell type specificity. Therefore, methods which allow unique cell types to be separately analyzed offer great insight into processes involving co-ordinated actions across multiple cell types.

Methods for isolating single cell types in plants: One approach to enable specific tissues to be studied is laser capture microdissection. A thin polymer film is attached to a tissue of interest and a microscope operator is able to use a laser to fuse the film to cells of the underlying tissue. By precise laser operation, it is possible to fuse specific clusters of cells to the film. When the film is removed, only fused cells will remain attached and these can then be analysed (Emmert-Buck *et al*., 1996).

Isolation of nuclei tagged in specific cell types (INTACT) is a method that specifically purifies the nuclei of individual cell types. This protocol utilizes plants transformed with a construct consisting of the first 111 amino acids of the *A. thaliana* RanGAP1 protein (which is sufficient for targeting to the nuclear envelope), a fluorophore and a peptide which is recognized by the *Escherichia coli* biotin ligase *BirA*. If the biotin ligase is also be expressed in parallel to the construct then biotin is produced, allowing recovery of transformed nuclei with streptavidin-coated magnetic beads. Analysis of nuclear RNA or DNA is then possible (Deal and Henikoff, 2011).

Whilst INTACT captures nuclear mRNAs, RNAs associated with cytoplasmic ribosomes can be profiled using translating ribosome affinity purification (TRAP). In this method, the ribosomes of a specific cell type are tagged with GFP. Cells can then be dissociated and lysed and GFP-tagged ribosomes can be purified by the use of anti-GFP antibodies attached to beads. mRNAs associated with ribosomes bound to the bead can then be recovered and sequenced (Heiman *et al*., 2014).

Recent advances in the resolution of RNAseq has facilitated transcriptomic profiling of single cells. The advent of Drop-seq allows this to be performed for thousands of cells in parallel. Primers containing a unique barcode are delivered on microparticles to cells suspended in droplets which are then lysed. Cellular mRNAs bind to the microparticle creating a complex

called a STAMP. STAMPS are then subject to reverse transcription, amplification and RNA sequencing. Barcode analysis then resolves the cell from which each profile originated (Macosko *et al*., 2015).

An alternative method to isolate single cells is fluorescence activated cell sorting (FACS). FACS is a technique that is used to segregate populations of cells based on their fluorescence characteristics. Cells in liquid suspension are injected into a stream of pressurized saline sheath fluid which is then vibrated and broken into droplets. Either prior to, or just after, droplets breaking off from the stream, suspended cells are interrogated by lasers and those which fluoresce within a user-determined set of parameters are deflected out of the stream by electrically charged plates and into a collection tube and are thus sorted.

Methodology of FACS: FACS facilitates the study of plant processes at the cell type-specific level and thus provides insights into plant biology that can be obscured at the whole organism or organ level, especially in the cases where genes show mixed patterns of induction and repression between cell types. FACS is particularly well suited to studying developmental processes within each layer of the plant root. Transgenic lines can be generated expressing a fluorophore, such as GFP, in a single cell type (Carter *et al*., 2013). A liquid suspension of plant cells can be achieved by a process called protoplast generation in which root or leaf material is treated with cellulose- and pectin-degrading enzymes that digest cell walls such that cells become protoplasts and no longer adhere to one another. The cell suspension can then be subject to FACS (Figure 1.5) and the purified cell type of interest can be subject to a range of downstream analysis (Birnbaum *et al*., 2005). Although FACS has predominantly been used to study the transcriptome of single cell types, studies characterising the metabolome (Moussaieff *et al*., 2013), proteome (Petricka *et al*., 2012) and chromatin structure (Frerichs *et al*., 2019) of a developing *A*. *thaliana* root demonstrate the feasibility of other applications.

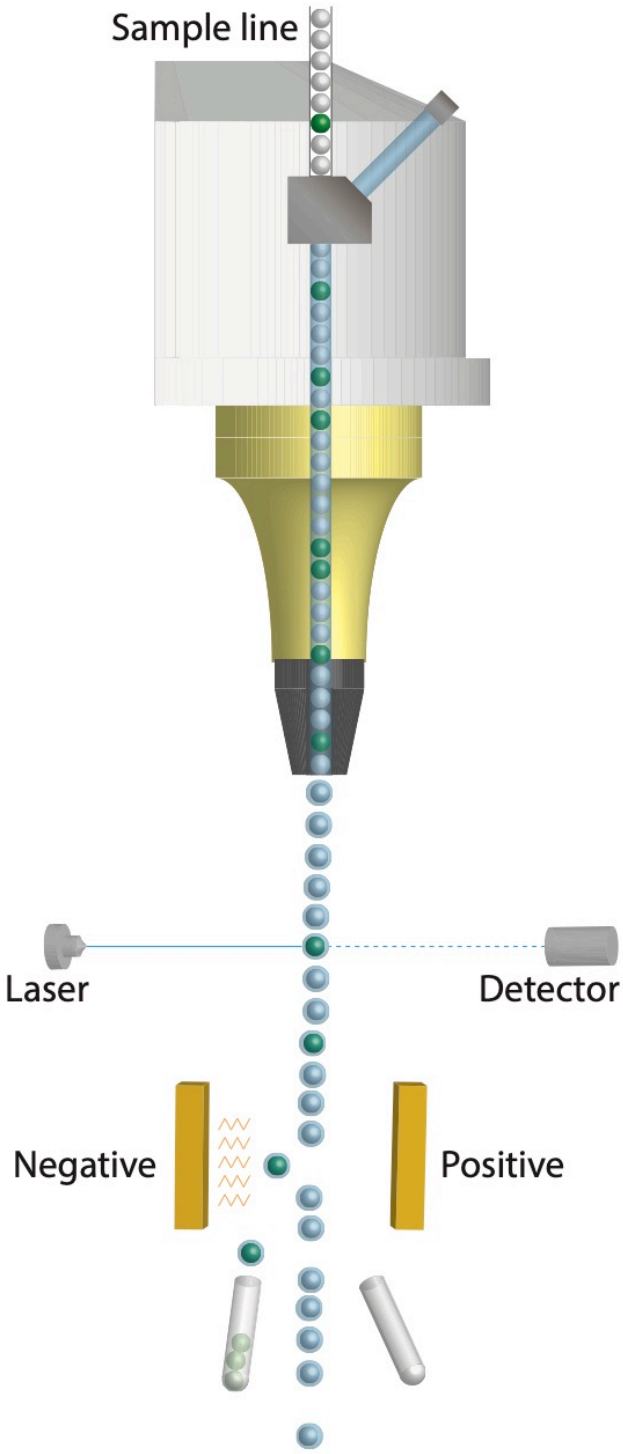

Figure 1.5: Fluorescence activated cell sorting of plant protoplasts. A cell sorter consists of a fluidics system that facilitates the flow of a saline sheath fluid through the machine. Pressurized sheath fluid travels through the machine to the nozzle. Separately, a sample containing particles or cells of interest (protoplasts in this case) is drawn up through the sample line to the nozzle. The sample is drawn at a slightly higher pressure, resulting in the sample being injected directly into the centre of the stream through a process

termed hydrodynamic focusing. As the stream exits the nozzle it is vibrated at high frequency, leading to formation of droplets from the stream. Droplets pass through a laser aligned to a detector behind the stream allowing the characteristics of particles/cells in each droplet to be determined (such as forward scatter, side scatter and fluorescence). If the fluorescence of an object fits a user-determined set of parameters of fluorescence intensity, termed 'gating', an electrical charge is applied to that droplet. Due to saline nature of the sheath fluid forming the outside of protoplast-containing droplets, this is sufficient to deflect the droplet when it passes the charged electrical plates. This deflection diverts the droplet into a collection tube. Upon conclusion of the sort, this receptacle will contain gated cells or particles of interest which can then be used for subsequent analysis.

FACS as a tool for studying plant roots: For the reasons explained above, FACS is an excellent tool for investigating root development. The first such published study in plants used *A*. *thaliana* transgenic lines with GFP expression localized to epidermis, cortex and endodermis, endodermis, stele and lateral root caps to isolate and transcriptomically profile the aforementioned cell types amongst different zones along the length of the root and thus at different developmental stages (Birnbaum *et al*., 2003). In addition to providing insight into key genes and hormone networks involved in root development, this also proved the viability of FACS as a tool to study root development at tissue specific level. It also identified that only a subset of genes (~250) consistently exhibit changes in expression in response to protoplast generation; these must be accounted for during such studies. This study has since been followed up with additional cell types and developmental zones to provide a higher resolution data set for the entirety of the plant root (Brady *et al*., 2007).

Since then, there are numerous examples of studies using FACS to better understand root biology. Some of these have focused on transcriptome responses to environmental changes in specific cell types and how this influences root development. The response of many root cell types to high salinity stress have been characterized (Dinneny *et al*., 2008; Geng *et al*., 2013) in addition to acidity and sulfur deficiency (Iyer-Pascuzzi *et al*., 2011)

and the transition from a deplete to replete nitrate environment (Alvarez *et al*., 2014; Gifford *et al*., 2008). Finally, the response of cortex and pericycle cell types to the immunogenic peptide flg22 and DAMP PEP1 has also been investigated (Rich *et al*., 2019). These studies found that although there is a degree of conservation of environmental responses between distinct cell types, the predominant component of these responses is cell type-specific.

Because of the applicability of FACS as a tool to study root development, studies have also focused on plant hormones, including the key regulator of development, auxin. Localized concentrations of auxin has been measured throughout the root (Petersson *et al*., 2009) and changes in gene expression in response to auxin subsequently identified (Bargmann *et al*., 2013b). Cells of the quiescent centre, where auxin levels were found to be highest and from where other root cell types arise, have also been isolated and transcriptomically profiled using FACS (Nawy *et al*., 2005). Finally, the role of *ACR4* (Arabidopsis CRINKLY4) in sorted pericycle cells has been probed, implicating this gene in regulating founder cell division during LR development (De Smet *et al*., 2008).

A novel application of FACS is transient assay reporting genome-wide effects of transcription factors (TARGET). In this approach, protoplasts are transfected with a plasmid carrying RFP as a fluorescent marker and a fusion protein of the TF interest with a glucocorticoid receptor (GR). By applying DEX, the TF can be moved to the nucleus on command therefore allowing the TF to interact with its targets. FACS can be used to isolate RFP+ (and therefore transfected with TF-GR fusion) which are then transcriptomically profiled to identify genes induced by the TF. By comparing protoplasts treated with or without the translation inhibitor CHX, direct and indirect targets of the TF can be discriminated against (Bargmann *et al*., 2013a). This approach has been used to identify targets of the transcription factor *HRS1* (Hypersensitivity to low P_i-elicited primary root shortening 1) in root protoplasts which is known to be important in regulating the response to nitrate (Medici *et al*., 2015).

Aims and rationale for work in the thesis:

Because LRs originate in the pericycle, it is likely that there are vast transcriptional differences between the pericycle and cell types following the application of a treatment that induces LR formation. To test this hypothesis, cells of the pericycle, as the site of LR emergence, and cells of the overlaying cortex were analyzed over time (Chapter 3: Regulation of gene expression by nitrogen and rhizobia underpinning lateral root emergence in *Arabidopsis thaliana*). In addition to monitoring untreated cells, the response to nitrate provision which is strongly linked to developmental changes in roots and the root architecture-modifying legume symbiont *Sinorhizobium meliloti* was also explored. The resulting dataset was both cell type-specific and longitudinal allowing these responses to be analyzed in an unprecedented level of detail.

The formation of root nodules of legumes represents another example of an adaptation of a plant root system to the environment. In this case, the plant must integrate both abiotic (nitrogen starvation) and biotic stimuli (rhizobia) to enable the outcome of nodule formation. The plant must also allow the rhizobia to enter, and colonize, the root without permitting this to be used as a route for pathogens to attack the plant. Suppression of defence responses is essential for nodulation to be maximized and it is likely that changes in gene expression are responsible for this transition from a state of defence to symbiosis by selectively dampening immune responses to in response to rhizobia. To identify these genes, the response to the model legume *Medicago truncatula* to the nitrogen fixing symbiont *Sinorhizobium medicae* and the pathogen *Ralstonia solanacearum* were analyzed in parallel to identify genes which are involved in the discrimination of symbionts and pathogens (Chapter 4: Molecular changes underpinning 'friend vs. foe' recognition in legume roots) by comparing symbiotic and defence responses.

Chapter 2: Methods

Plant material and growth conditions:

***Arabidopsis thaliana*:** *A. thaliana* GFP-expressing lines in the Col-0 background to mark the cortex (ProCo2:YFPH2B line (Heidstra *et al*., 2004)) or pericycle (E3754 enhancer trap GFP line (Gifford *et al*., 2008)) were used. Approximately 400 seeds per plate were surface-sterilized in 100% ethanol for 5 minutes, 7 % sodium hypochlorite solution for 5 minutes and washed 5 times in sterile water. Sterilized seeds were then sown on 0.7 cm strips of autoclave-sterilized brown growth pouch paper (CYGTM Germination Pouch, West St. Paul, MN, USA) on 0.8% agar plates containing a basal 1x Murashige and Skoog (MS) growth medium (Sigma-Aldrich M0529), supplemented with sucrose (30 mM), $CaCl_2$ (1.5 mM), $MgSO_4$ (0.75 mM), KH_2PO_4 (0.625 mM) and 0.3 mM NH_4NO_3; adjusted to pH 5.7 with KOH. Plates were sealed with microporous tape, seeds stratified for 2 days at 4°C in the dark, then plates placed in opaque (black polythene) covers (Bagman of Cantley, Lingwood, Norfolk, UK) and grown in a Sanyo growth chamber (MLR-351H, Sanyo, E&E Europe BV, Loughborough, UK) with a 16/8 h photoperiod at 50 mmol m^{-2} s^{-1} at a constant 22 °C.

***Medicago truncatula*:** Seeds of *M. truncatula* ecotype A17 were scarified in 95% sulfuric acid for 25-30 mins until dark spots appeared on all seeds. Sulfuric acid was removed and replaced with 10 ml of water followed by 2 more subsequent wash steps. Seeds were then surface sterilized in 7 % sodium hypochlorite solution for 5 minutes and washed 8 times in sterile water. Seeds were then sown onto plates containing 1.5% water agar. 1 drop of sterile water was added to each seed and left until completely soaked up. This step was repeated a further 2 times. Two growth pouches (CYG™ Germination Pouch, Mega International, Newport, MN, USA) were added to the lid of the plate and soaked with sterile water before the plates were sealed with microporous tape. Plates were kept at 4 °C for seed stratification for 2 days and then transferred to a growth chamber at 24 °C for 2 days in the dark.

Seedlings with a radial length of >2 cm were transferred to pots containing perlite with a ~ 1 cm layer of vermiculite on top. Pots were watered with nutrient solution (1 mM $CaCl_2$, 1 mM KH_2PO_4, 75 µM FeNaEDTA, 1mM $MgSO_4$, 0.25 mM K_2SO_4, 6 µM $MnSO_4$, 20 µM H_3BO_3, 1 µM $ZnSO_4$, 0.5 µM $CuSO_4$, 50 nM $CoSO_4$, 0.1 µM Na_2MoO_4, adjusted to pH 6.5-6.8 with KOH) with (replete nitrate mock, ralstonia-treated or flg22-treated) or without (deplete nitrate mock or rhizobia-treated) 15 mM ammonium nitrate. Pots were also incubated at 24 ˚C with a 16/8 hour photoperiod at 50 mmol m^{-2} s^{-1}. After 1 day in pots, seedlings were mock-treated, treated with *Ralstonia solanacearum*, treated with *Sinorhizobium meliloti* or treated with flg22 (see treatment regimes below) at dawn. Whole root systems were sampled immediately prior to treatment and 30 minutes, 1 hour, 3 hours, 6 hours, 12 hours, 1 day, 2 days, 4 days, 6 days, 9 days and 12 days after treatment and immediately flash frozen and stored at -80 °C.

Plant treatments:

***A. thaliana* treatment with nitrogen or *S. meliloti*:** After 9 days in the growth chamber, treatments were carried out at dawn (= time 0). To prepare the *Sinorhizobium meliloti* solution, 50 mL of TY/Ca^{2+} medium (5 g/L Bacto-tryptone, 3 g/L Yeast extract, 6 mM $CaCl_2$) (Journet *et al*., 2001) was inoculated with a single colony of *S. meliloti* and incubated for 26 h at 28 °C and 220 rpm to an OD_{600} of 1–2. Cells were harvested by centrifugation (4000 rpm, 10 min, 4°C) and re-suspended in water to a final OD_{600} of 0.01. To carry out the treatments, seedling strips were removed from plates, then briefly immersed (10 seconds) in liquid deplete N medium (as plates but with no agar) and placed on a fresh plate (Untreated (U), or briefly immersed in liquid replete N medium (as plates but with no agar and with 5 mM NH_4NO_3) and placed on a fresh 5 mM NH_4NO_3 plate (N treatment), or briefly immersed (10 seconds) in a solution of *S. meliloti* and placed on a fresh deplete NH_4NO_3 plate. Plates were then re-sealed with microporous tape, replaced in opaque (black polythene) covers, and returned to the Sanyo growth chamber. At hourly time points 0 (immediately before transfer, at dawn), 1, 2, 4, 6, 8, 10, 12, 14, 16, 20, 24, 36 and 48 h after transfer, whole roots were harvested for protoplast

generation and FACS by cutting roots just below the growth pouch paper strip. The whole experiment was carried out in at least biological triplicate with seedlings for each biological replicate grown independently (in a non-overlapping time period). Each replicate sample (for each time point in each time series) consisted of the pooled harvested roots from ~200 GFP-expressing seedlings (destructive sampling) grown on the same plate. Each replicate sample was considered independently within subsequent analysis.

M. truncatula **treatment with flg22, *R. solanacearum* or *S. medicae*:** *S. medicae* strain WSM419 (kindly provided by Jason Terpolilli, Murdoch University) was grown on Ty/Ca^{2+} plates (20 mM HEPES, 0.5% tryptone, 0.3% yeast extract, 6 mM $CaCl_2$ and 1.5% bactoagar adjusted to pH 6.8-7.0 with KOH and supplemented with 50 µg/ml chloramphenicol) for 2 days at 28 °C. Liquid cultures were obtained by inoculating a single colony into 10 ml liquid Ty/Ca^{2+} medium (as above including chloramphenicol, minus agar) for 2 days in a shaking incubator at 28 °C and 220 rpm. Cultures were then centrifuged for 10 mins at 2880 rcf and 4 °C before the pellet was re-suspended in sterile water and diluted to a final OD_{600} of 0.05.

R. solanacearum strain GMI1000 was grown on BGT media (1% bactopeptone, 0.5% glucose, 0.1% yeast extract, 0.1% casamino acids, 0.0005% tetrazolium chloride, 1.5% agar and adjusted to pH 7.2) for 2 days at 28 °C (Boucher *et al*., 1985). Liquid cultures were obtained by inoculating a single red colony (red colour indicating virulence (Boucher *et al*., 1985)) into liquid BGT media (as above, minus agar) for 2 days in a shaking incubator at 28 °C for 2 days. Cultures were then centrifuged for 10 mins at 2880 rcf and 4 °C before the pellet was re-suspended in sterile water and diluted to a final OD_{600} of 0.8.

For flg22 assays, plants were treated with a 100 nM flg22 solution. For all treatments, plants were treated by dispensing a volume of 250 µl flg22/inoculant on the perlite in extremely close proximity to the root or directly onto the root during flg22 treatment.

Cell sorting:

Protoplast generation and Fluorescence Activated Cell Sorting (FACS): Harvested roots were bundled and cut into lengths of approximately 2–3 mm into a 70 µm cell strainer in a small petri dish containing protoplast-generating solution (600 mM mannitol, 2 mM MES hydrate, 10 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$ and 0.1% BSA, pH adjusted to 5.7 with Tris HCl) with 1.5% cellulase R-10 (Phytotechlab), 1.2% cellulase RS (Sigma), 0.2% macerozyme R-10 (Phytotechlab), 0.12% pectinase (Phytotechlab)). The petri dishes were placed on an orbital shaker at 200 rpm for 1 hour then cells harvested by centrifugation (5 mins at 1200 rpm at room temperature), resuspended in 500 µL protoplast-generating solution lacking enzymes, then filtered through a 40 µm cell strainer to break up large clumps of protoplasts.

Protoplasts were sorted using a BD Influx cell sorter (BD Biosciences) fitted with a 100 µm nozzle, using FACS-Flow (BD Biosciences) as sheath fluid. Pressure was maintained at 20 psi (sheath) and 21–21.5 psi (sample), drop frequency was set to 39.5 kHZ, which yielded an event rate of <4000; these are optimal settings on a BD influx cell sorter for this type of protoplasts (Gifford *et al*., 2008). To optimize alignment of lasers and detectors, Calibrite Beads (BD Biosciences) suspended in FACS-Flow were used, and to optimize sort settings, BD Accudrop Fluorescent Beads (BD Biosciences) suspended in FACS-Flow were used. GFP positive protoplasts were identified using a 488 nm argon laser by plotting the output from the 580 nm/30 nm, bandpass filter (orange) vs. the 530 nm/40 nm bandpass filter (green). GFP/YFP positive protoplasts were in the high 530 nm/low 580 nm population, with non-GFP protoplasts in the low 530 nm/low 580 nm population and dead/dying protoplasts and debris in the high 580 nm population (Gronlund *et al*., 2012). At least 10,000 GFP positive protoplasts were sorted using methods shown previously to preserve endogenous gene expression levels (Birnbaum *et al*., 2003; Gifford *et al*., 2008). Sorted protoplasts were directly sorted into Qiagen RLT lysis buffer containing 1% (v:v) β- mercaptoethanol, mixed and immediately frozen at –80°C for RNA extraction (methods according to (Gifford *et al*., 2008)).

Molecular methods:

RNA extraction: For *A. thaliana* samples, RNA was extracted from sorted cells using the Qiagen RNeasy RNA Kit, and DNase treatment carried out using the Qiagen DNase Kit. The quantity and quality of RNA was checked with a Bioanalyzer 2100 RNA 6000 Pico Total RNA Kit (Agilent Technologies). For *M. truncatula* samples, frozen root material was homogenized using a mortar and pestle before extraction of RNA using a Monarch RNA MiniPrep Kit (New England Biolabs) according to the manufacturer's protocol, including an on-column DNase treatment. Absence of contaminating genomic DNA was confirmed by PCR using primers directed against the gene Medtr3g091400. RNA was stored at -80 ˚C.

cDNA synthesis: For *A. thaliana* samples, cDNA was amplified from RNA using ½ reactions of the Ovation Pico WTA System V2 Kit (NuGEN Technologies Inc., San Carlos, CA, USA) according to the 'quick' protocol. Post-amplification purification of cDNA was performed using the QIAquick PCR Purification Kit (Qiagen). For *M. truncatula* samples, reverse transcription was performed using ProtoScript II cDNA synthesis kit (New England Biolabs) according to the manufacturer's instructions using 300 ng of total RNA. cDNA was stored at -20 ˚C.

qPCR: Primers for qPCR were designed with the following parameters: primer melting temperature between 59 and 61˚C (60˚C optimum), product size 50-250 BP and a GC content of between 45-55% (50% optimum). qPCR was performed in 96 well plates using SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich). Each sample for analysis was loaded into 3 separate wells with 9 µl final reaction volume per well. Thermocycler conditions are displayed in Table 2.1. Expression of genes of interest was normalized against the reference genes Medtr3g091400, Medtr7g116940 and Medtr7g089120 (Kakar *et al*., 2008) and differential expression of treated samples relative to mock samples was calculated using the ΔC_t method, a derivation of the $\Delta\Delta C_t$ (Livak and Schmittgen, 2001). Significant differences between treated and mock samples were calculated using the Student's t-test.

Temperature/Time (mins)					
95 °C 10:00	95 °C 0:30	60 ˚C 1:00	95 °C 1:00	60 ˚C 0:30	95 ˚C
	X 40				

Table 2.1: Thermocycler conditions for qPCR.

Microarray analysis: 0.5 µg NuGEN-amplified cDNA was labeled using the NimbleGen One-Color Cy3 Labeling Kit, and 4 µg of this was hybridized using the GeneChip Hybridization Kit on NimbleGen *A. thaliana* 12 x 135k probe microarrays designed for the full TAIR-10 annotation *A. thaliana* genome (design OID 37507; this can be found by searching "GPL18735" in the Gene Expression Omnibus NCBI store (Roche Applied Science, Upper Bavaria, Germany)). OID has 131,524 probes designed against 27,143 genes in the *A. thaliana* genome; 3,675 genes had multiple probes along their length to determine expression of 31,524 gene transcripts (an average of 2.19 and up to 4 gene models each for these genes). Arrays were imaged using an MS200 microarray scanner using only the 480 nm laser and the autogain feature of the NimbleScan software; all according to the manufacturer's instructions. Grids were aligned automatically then manually verified. Raw probe (xys) and gene level unnormalized data were obtained using NimbleScan. Expression data from the arrays can be found by searching for "GSE91379" in the Gene Expression Omnibus NCBI store.

RNAseq processing: Extracted RNA from three replicates of samples from roots treated with *R. solanacearum* and *S. medicae* 3 and 12 hours after treatment were sent for sequencing using an Illumina NovoSeq 6000 (Novogene UK). Replete nitrate mock and deplete nitrate mock T=0 samples were used respectively as references for *R. solanacearum* and *S. meliloti* samples. Selection of mRNAs was performed on the basis of poly(A)-purification. Sequencing was paired end and unstranded with a read length of 150 BP and a depth of sequencing of at least 20 million reads/sample. Raw data was obtained in the form of .fq.gz files for each pair of each sample.

Bioinformatics and statistics:

Normalization and quality assessment of microarray data: The xys (NimbleGen) files, which store array coordinates and observed intensities, were imported with the Bioconductor 'oligo' package (Carvalho and Irizarry, 2010) using the annotation package pd.120110.athal.mg.expr installed through the pdInfoBuilder package. The RMA algorithm, that performs background subtraction, quantile normalization and summarization via median polish, was applied to the raw data of expression arrays to obtain the $\log_2$ normalized gene expression levels. All arrays for pericycle and cortex were normalized together, and arrays were assessed for quality implemented by the Bioconductor package arrayQualityMetrics (Kauffmann *et al*., 2009). An object of class ExpressionSet, which was generated by the oligo package, was inputted to the arrayQualityMetrics package then a utility report created. To generate a robust total dataset, outlier arrays were removed based on the between array distances (using the sum of distances Sa as the quality metric; with a cutoff of 337), boxplot (using Kolmogorov-Smirnov statistic Ka as the quality metric; with a cutoff of 0.063) and MA-plot (using Hoeffding's statistic Da as the quality metric; with a cutoff of 0.15).

Determination and clustering of differentially expressed genes: Differentially expressed (DE) genes within each time series were obtained using the software BATS (Angelini *et al*., 2008) and robustness of assignment confirmed using the independent method of gradient analysis (Breeze *et al*., 2011). The BATS input file for each time series contains the rescaled $\log_2$ gene expression values such that the vector of $\log_2$ expression values of each gene has mean zero and variance one. Genes were considered to be DE if their Bayes Factors in the BATS output file were less than a threshold ($\log_2$ Bayes Factor >3), which was determined by the histogram of $\log^{10}$ of Bayes Factors and the regression plots of gene expression levels from BATS. Differentially expressed genes between treatments in each cell type were obtained using Gaussian process modelling implemented in the software GP2S (Stegle *et al*., 2010). In the GP2S input file, the $\log_2$ expression levels of each gene for the two time series were rescaled such that their mean was zero and their variance

was one. GP2S assigned each gene a Bayes Factor that equals the difference between the likelihood of the expression level of a gene in two treatments being sampled from different Gaussian processes and that from the same Gaussian process.

Immediate responding genes were defined as those with $|x_1-x_0| > \varphi_1 \times sd(x_t)$ for gene transcription time series x_t, i.e., the absolute value of the change in the first time point being larger than φ_1 times the time series standard deviation. Furthermore, fast and downstream responding genes were defined as genes that were not immediate responders, but had a dramatic change within 2 or 6 h of treatment ($sd(x_{0..2}) > \varphi_2 \times sd(x_t)$, $sd(x_{0..6}) > \varphi_6 \times sd(x_t)$). The remaining genes not in any of these three classes were designated late. Clusters were further defined as up-regulated or down-regulated if their expression increased or decreased at this point of differential expression. The thresholds φ_1, φ_2, and φ_6 were chosen, then clusters were visually inspected to confirm separation of the clusters into the appropriate classes.

Hierarchical clustering of the differentially expressed genes was performed using SplineCluster (Heard *et al.*, 2005). The mean expression levels of biological replicates of a DE gene at each time point in a time series was used as input. A reallocation function was implemented to reallocate outliers of each cluster into other more appropriate clusters at each agglomerative step. A prior precision value was determined after trying different values and comparing their effects on clusters. Standard error of the mean was plotted on all graphs of gene expression.

Gene regulatory network inference and analysis: The Bioconductor package GRENITS (Wang *et al.*, 2014) was used for gene regulatory network inference using the mean expression levels of biological replicates (as outlined above) for each DE gene at each time point from 0 to 16 hours. Genes that were DE within time series (as calculated using BATS) were used as input, then genes DE when comparing treatment and untreated control samples (GP2S) were identified within those networks. GRENITS gave the posterior probability of each directed link between two genes. A link probability threshold was chosen based on the confidence required for assigning links, then a link

in the gene regulatory network was assigned if its posterior probability was greater than or equal to the link probability threshold. Predicted edges were compared with the interactions found by O'Malley *et al*. using ampDAP-Seq and DAP-Seq (O'Malley *et al*., 2016) where possible.

RNAseq data processing: Initial quality control of raw reads was performed using FastQC. Raw reads were then subject to trimming and adaptor removal using Trimmomatic v0.33 (Bolger *et al*., 2014) with the following parameters; LEADING = 20, TRAILING = 20, SLIDINGWINDOW = 4:20, MINLENGTH = 75 and using PHRED33 quality scores. FastQC was again used for quality control of the clean reads.

Alignment of trimmed reads was performed using STAR (Dobin *et al*., 2013) using only paired reads. The index files for STAR were generated with files of the *M. truncatula* genome (Tang *et al*., 2014) and the *M*. *truncatula* genome annotation 4.0v2 with the sjdbOverhang parameter set to 149 (read length -1).

Raw counts data for 50,443 genes in the *M. truncatula* genome was obtained from the alignment files using LiBiNorm (Dyer *et al*., 2019) in htseq-compatible mode again using the genome annotation *Mt*4.0v2 with intersection mode set to strict. Differential expression analysis was conducted using DESeq2 (Love *et al*., 2014). Genes were considered DE if they exhibited a $\log_2$ fold change not between 2 and -2 and a Bonferroni-corrected P-value <0.05. Lists of DE genes generated from DESeq2 were queried for overrepresented processes using AgriGOv2 (Tian *et al*., 2017). Locus ID V4.0 (JCVI) was used as a background and GO terms were considered significantly overrepresented a Bonferroni-adjusted P-value <0.05 was returned.

Chapter 3: Regulation of gene expression by nitrogen and rhizobia underpinning lateral root emergence in *Arabidopsis thaliana*

Introduction:

Control of nitrate metabolism in the model plant *A. thaliana*:

Nitrogen is essential for plant growth due to its incorporation in amino acids and nucleotides. Most plant species rely entirely on deriving nitrogen required for growth from nitrate or ammonia from the soil and thus its availability is a common limiting factor for growth (reviewed by Crawford, 1995). Plants have adapted elaborate strategies to regulate their uptake and storage of nitrate and in the longer term can alter their root architecture to adapt to levels of nitrate availability (Krouk *et al*., 2010a).

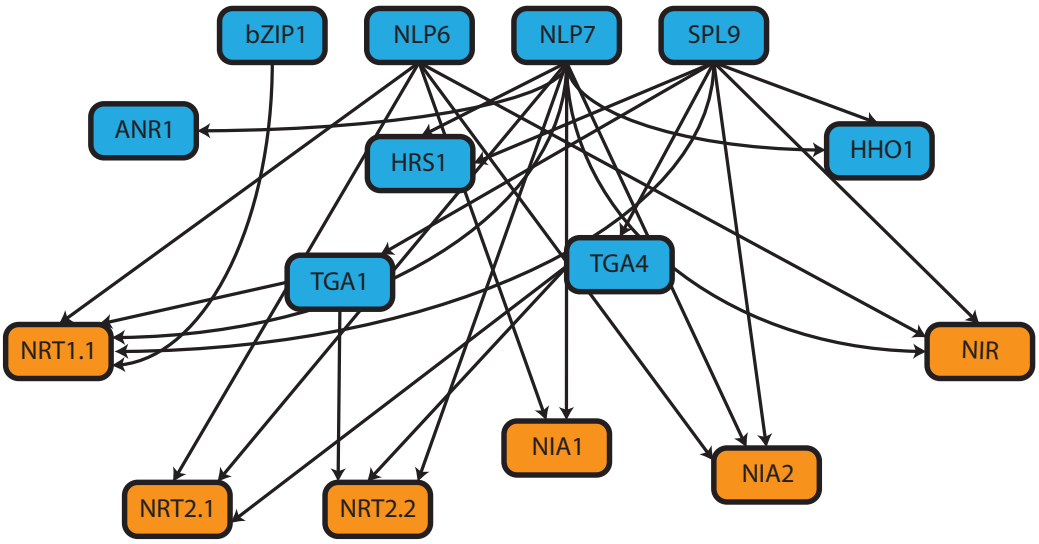

Figure 3.1: Regulation of select genes during nitrate responses in *A. thaliana*. Select transcription factors associated with the primary nitrate response of *A. thaliana* (shaded blue) and genes associated with the transport and metabolism of nitrate (shaded orange) are displayed as nodes. Edges denote reported regulatory interactions with arrowheads indicating positive regulation. Adapted from Vidal *et al* (Vidal *et al*., 2015).

Nitrate transporter systems in *A. thaliana*: Upon nitrate provision, its uptake into plant roots is mediated by nitrate transporters. There are two mechanisms of nitrate uptake mediated by different classes of nitrate transporters in plants; high affinity (HATS) and low affinity (LATS) transport systems with the former further divided along the lines of constitutively expressed or inducible transporters. In *A. thaliana*, there are four characterized families of nitrate transporters; *NRT1*/*PTR*, *NRT2*, *CLC* and *SLAC1*/*SLAH*. Although collectively these families encompass 73 genes, not all are functional in nitrate transport. Very little is known about the latter two families and they are not thought to be involved in uptake of nitrate in the root (Krapp *et al*., 2014).

The *NRT1*/*PTR* family consists of 53 proteins involved in the transport of various molecules (Krapp *et al*., 2014). Nitrate transporters belonging to this family are considered low affinity nitrate transporters with the exception of *NRT1.1* which is capable of transitioning to high affinity nitrate transport upon phosphorylation of a cytoplasmic threonine residue (Liu and Tsay, 2003). *NRT1.2* is also implicated in nitrate transport into the root (Huang *et al*., 1999). Although it has not been observed to be involved in nitrate uptake, knockout of *NRT1.5* results in a reduction (although not elimination) of root-to-shoot nitrate transport, suggesting a role in this process. As this protein is expressed in the pericycle, it likely functions, possibly alongside *NRT1.8*, in loading nitrate into the xylem for transport to aerial tissues (Chen *et al*., 2012; Lin *et al*., 2008). *NRT1.6* is expressed in seeds which show reduced nitrate accumulation in its absence. It is therefore likely that this protein is responsible for loading nitrate from the plant into developing seeds (Almagro *et al*., 2008). *NRT1.7* is expressed in leaf tissues and seems to be involved in loading nitrate into the phloem of old leaves for reallocation to younger leaves (Fan *et al*., 2009).

The *NRT2* family consists of seven genes and all have been designated as high affinity transporters specifically for nitrate (Orsel *et al*., 2002). All *NRT2s* except *NRT2.7* have been demonstrated to interact with an additional protein, *NAR2.1* and this interaction enhances their function as nitrate transporters in some cases (Kotur *et al*., 2012). *NRT2.4* has a very high affinity for nitrate, suggesting it is responsible for nitrate uptake under conditions of enduring nitrate starvation (Kiba *et al*., 2012). However, when nitrate is available, *NRT2.1* is responsible for the bulk of its transport into the root with

a more minor role for *NRT2.2* (Li *et al*., 2007). Like *NRT1.1*, *NRT2.1* also seems to function as a nitrate sensor and represses LR emergence during nitrate starvation but a mechanism has not yet been described by which this is achieved (Little *et al*., 2005).

The primary nitrate response: When nitrate becomes available, it induces a massive reprogramming of transcription termed the primary nitrate response (PNR). This transcriptional response still occurs in the absence of nitrate reductase activity or the presence of protein synthesis inhibitors which suggests that it is directly induced by nitrate as opposed to any of its derivatives (Medici and Krouk, 2014). This pathway has become much better understood (Figure 3.1) since the discovery of the nitrate responsive element (NRE) in the promoter of the nitrate response marker gene *NIR1* which has accelerated the discovery of transcription factors involved in regulating this response (Konishi and Yanagisawa, 2010).

One key component found to be involved in the PNR is *SPL9* (SQUAMOSA PROMOTER BINDING-LIKE 9). This gene was identified by a time series approach monitoring transcriptome variation during very early nitrate responses which was in turn used to infer networks (Krouk *et al*., 2010c). *SLP9* was found to be expressed within 10 minutes of nitrate provision, and amongst its targets there were known markers of the nitrate response. Although *SLP9* mutants do not show any obvious phenotypes, likely due to redundancy in the N-response network, overexpression resulted in increased expression of the targets it was predicted to regulate (Krouk *et al*., 2010c).

The TF *NLP7* (Nodule Inception-like protein 7) is also a major component of the PNR. *nlp7* mutants are deficient in nitrate assimilation regardless of availability and as such produce phenotypes resembling nitrate-starved plants (Castaings *et al*., 2009). Conversely, its overexpression results in increased nitrate uptake even under nitrate-limiting conditions (Yu *et al*., 2016). *NLP7* has been demonstrated to very rapidly control hundreds of genes upon nitrate provision, amongst which nitrate-responsive genes are highly over-represented. *NLP7* is not transcriptionally induced by nitrate but rather seems to be retained in the nucleus in its presence (Marchive *et al*., 2013).

Collectively, this demonstrates *NLP7* is critical in co-ordinating the PNR. Another related protein, *NLP6*, is also thought to be important as it has been demonstrated to bind to the NRE and its constitutive suppression prevents regulation of normally nitrate-responsive genes (Konishi and Yanagisawa, 2013).

The *ANR1* TF is another key transcriptional regulator of the PNR. Because its expression is increased in *NLP7* overexpressing mutants, it is thought to function downstream of *NLP7* (Yu *et al*., 2016). Whilst little is known about the targets of *ANR1*, its importance is seen in the failure of plants to produce LRs in response to nitrate in its absence (Gan *et al*., 2005; Remans *et al*., 2006; Zhang and Forde, 1998).

HRS1 and its homolog *HHO1* (HRS1 homolog 1) (Marchive *et al*., 2013), which are implicated in nitrate responses, are also downstream of *NLP7* (Alvarez *et al*., 2014). These genes are thought to be partially redundant and are responsible for suppressing primary root elongation during phosphorus starvation only in the presence of nitrate. Additionally, genes regulated by *HRS1* have been identified using the TARGET approach described in Chapter 1 (Medici *et al*., 2015). *HRS1* is nitrate-inducible and represses many genes associated with nitrate starvation, such as the previously described nitrate transporter *NRT2.4*. This is consistent with a possible role in signalling recovery from nitrate starvation (Kiba *et al*., 2018).

The TFs *TGA4* and *TGA1* were identified as components of the PNR by a data-mining approach. Both are transcriptionally induced by nitrate and double mutants for both genes produce a similar LR phenotype to *ANR1* mutants. Amongst their inferred targets are the nitrate transporters *NRT2.1* and *NRT2.2* (Alvarez *et al*., 2014).

Mutations in the TF *bZIP1* have been demonstrated to perturb nitrate signalling in a number of studies including (Obertello *et al*., 2010). Subsequent experiments combining approaches to identify genes induced by its transient induction and the DNA sequences that it binds through ChIP-seq have shown that bZIP1 (basic Leucine Zipper domain 1) is also implicated in the PNR due to overlap between its putative targets and nitrate responsive genes (Para *et al*., 2014).

Aside from TFs, other genes found to be important in the PNR are responsible for processing nitrate itself. The nitrate reductase enzymes *NIA1* and *NIA2* are responsible for converting nitrate to nitrite which is subsequently converted to ammonia by the nitrite reductase *NIR1*. These genes are commonly used as markers for the PNR and are regulated by most of the TFs described previously (Yanagisawa, 2014). Additionally, nitrate transporters involved in nitrate influx are themselves up-regulated, likely as a positive feedback mechanism to increase nitrate uptake. These include *NRT1.1* (which is essential for the activation of the PNR in the first place), *NRT2.1* and *NRT2.2* (Vidal *et al*., 2015).

Mutants affecting root architecture: One common reverse-genetics approach to identify genes involved in the regulation of root development is to study the root architecture phenotypes of mutants that show increased or reduced expression of a particular gene. Two of the most significant mutants found so far to affect *A. thaliana* root organisation are *scarecrow* (*scr*) and *short-root* (*shr*). Both of these genes are involved in regulating asymmetric cell divisions during root development and plants show dramatically altered root patterning in their absence. Both *shr* and *scr* mutants only form a single layer of cells between the epidermis and cortex. This mutant layer in *shr* does not express endodermal identity markers, implicating *SHR* in specifying cells of the endodermis whilst the corresponding layer in *scr* mutants exhibits cortical and endodermal markers, suggesting *SCR* functions specifically during the division of these layers (Di Laurenzio *et al*., 1996; Helariutta *et al*., 2000). Although *shr* mutants are able to initiate LR primordia, abnormal patterning is also present in lateral roots in addition to defects in elongation (Lucas *et al*., 2011).

Due to its role as a positive regulator of LR patterning, there are numerous mutants in the auxin pathway which display altered root architecture. The *Auxin/Indole-3-Acetic Acid (Aux/IAA)* family of proteins act as be negative regulators of auxin signalling. Gain of function mutation of *IAA3* (Tian and Reed, 1999), *IAA14* (Fukaki *et al*., 2002), *IAA19* (Tatematsu *et al*., 2004) and *IAA28* (Rogg *et al*., 2001) all result in plants that exhibit stunted growth and are almost completely unable to form lateral roots. Expression of

GATA23 rescues this phenotype, suggesting that *GATA23* acts antagonistically of *IAA28* (De Rybel *et al*., 2010). The TFs *ARF7* and *ARF19* and their downstream targets *LBD16* and *LBD29* are positive regulators of LR formation. Thus, *arf7arf19* double mutants are defective in lateral root formation and *LBD16*/*LBD19* overexpression partially rescues this phenotype (De Rybel *et al*., 2010; Okushima *et al*., 2007). LR outgrowth regulator *MYB93* mutants exhibit increased LR density whilst the opposite phenotype is observed in overexpressing lines (Gibbs *et al*., 2014).

In addition to their role in transducing defence responses in plants, there is also strong evidence that MAPK cascades are crucial in developmental responses. Mutants affected in expression of *MEKK1*, a kinase upstream of MAPK4, show severe dwarfism and late lethality (Ichimura *et al*., 2006). *mpk4* mutants also exhibit a similar phenotype although the seedling lethality was temperature dependent rather than constitutively lethal (Gao *et al*., 2008). MAPK3 and *MAPK6* are also important in plant development, evidenced by the fact that *mpk3*-*mpk6* double exhibit embryonic lethality (Wang *et al*., 2007).

As visited previously, transporters involved in root nitrate uptake and some of their downstream targets also have significant effects on root architecture. In addition to its role in nitrate uptake, *NRT1.1* is also important in sensing and signalling of the nitrate status of the plant as it facilitates the release of auxin from LR primordia, thus preventing their emergence. Because the auxin transport activity of NRT1.1 is repressed by nitrate, the protein inhibits LR emergence only in deplete nitrate conditions (Krouk *et al*., 2010b). Mutation of genes that function downstream of *NRT1.1* result in plants with altered root architecture. *ANR1* mutants fail to proliferate LRs in response to nitrate (Zhang and Forde, 1998) whilst *NLP7* mutants have increased LR density (Castaings *et al*., 2009). The *lateral root initiation 1 A. thaliana* mutant (*lin1*)/*NRT2.1* was found to constitutively form LRs even in the presence of high sucrose/low nitrate conditions which normally strongly repress outgrowth (Little *et al*., 2005; Malamy and Ryan, 2001).

Aims: Previous work in the Gifford lab had been carried out to profile untreated, nitrate-treated or *S. meliloti*-treated *A. thaliana* roots at a range of

timepoints up to 2 days after treatment as both of these treatments alter root architecture. These roots were subject to protoplast generation and FACS to isolate cells of the pericycle and cortex and subsequent transcriptome analysis using microarray profiling. This chapter sets out the data analysis conducted in order to identify regulators of root architecture remodelling, and contributing to the publication of this work in (published in Walker *et al*., 2017).

Results and Discussion:

Cells of the cortex and pericycle exhibit highly specific and dynamic gene expression through time: To identify transcriptional changes during, and therefore possibly underpinning, root architecture reshaping, cells of the pericycle (as the site of lateral root emergence) and cortex (an overlaying cell type) were sampled and transcriptomically profiled. This was perormed using protoplast generation followed by FACS of cell-type specific *A. thaliana* seedlings sampled over a range of timepoints from dawn (0 hours) up to 48 hours after nitrogen treatment or rhizobial inoculation (Figure 3.2A-E, see methods). Microarrays comprising probes designed against 31,524 transcripts of the *A. thaliana* TAIR10 annotation transcriptome were used. Untreated roots, roots treated with a replete level of ammonium nitrate (5mM NH_4NO_3) and roots inoculated with the bacterium *S. meliloti* (rhizobia) were sampled. *S. meliloti* is an organism normally associated with nitrogen fixing symbiosis in legumes, but it also induces root architecture changes in *A. thaliana* (as shown in (Walker *et al*., 2017)). Following quality control, a final set containing 236 comparable quality arrays (126 for pericycle and 110 for cortex) over the 6 treatments and 14 time points (average replication of 2.95 samples per time point) with at least 2 replicates per time point was used for analysis. Transcriptome data was normalized and then Bayesian Analysis of Time Series (BATS) (Angelini *et al*., 2008) was used to identify transcripts that exhibit differential expression (DE) during each experiment (over time within each time series). For subsequent experiments (nitrogen and rhizobia responses, discussed later) DE genes were computed both within the rhizobia/replete N experiment and relative to the baseline established in the PU and CU experiments (see methods) using Gaussian Process 2 Sample

(GP2S) (Stegle *et al*., 2010). This dual analysis was necessary to differentiate between genes that were DE during the experiment because of circadian effects (e.g. clock-related genes) and those that are actually DE *between* experiments (i.e. nitrogen or rhizobia responsive).

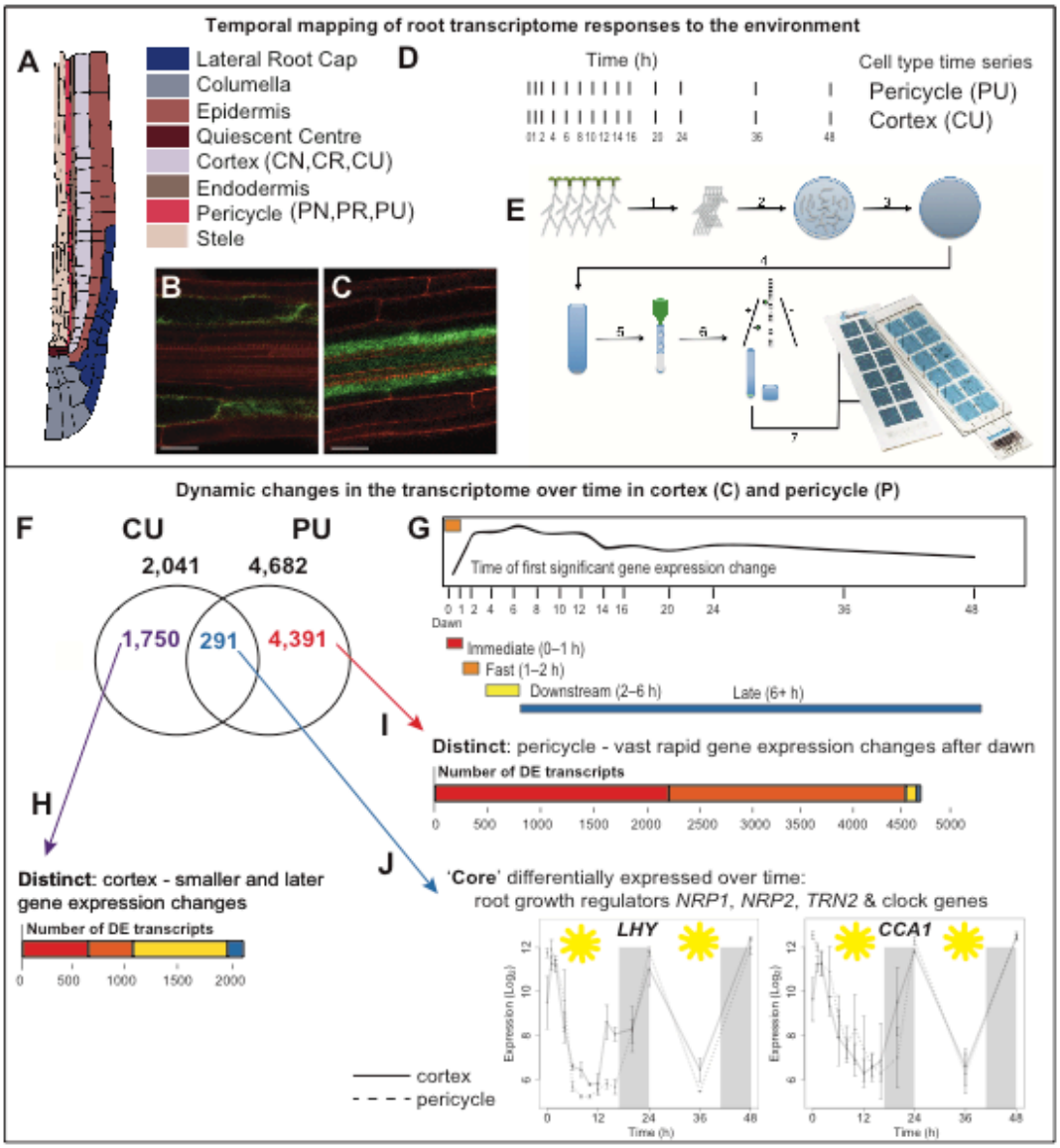

Figure 3.2. Overview of experimental design. (A) Diagrammatic representation of an *A. thaliana* root with different cell types labelled. (B-C) Confocal microscopy images of *A. thaliana* root where marker lines express GFP in the cortex (B) or the pericycle (C). Scale bars represent 25 μm. (D) Timepoints used for sampling in the study. (E) Diagrammatic representation of sampling workflow including protoplast generation, FACS and microarray analysis. (F) Venn diagram showing the differentially expressed (DE)

transcripts within the untreated time series in the cortex (CU) and pericycle (PU); (G, Key to timings assigned to genes through clustering; red – immediate (0-1 h), orange – fast (1-2 h), yellow – downstream (2-6 h) and blue late (6+ h). (H-I) Numbers of differentially expressed transcripts over time in the cortex (H) and the pericycle (I) with timing of first significant change in expression denoted by the colour coding set out in (G). (J) 'Core' regulated transcripts including $\log_2$ expression plots of the clock genes *LHY* and *CCA1* showing the similarity of their expression between pericycle and cortex cells. Error bars represent standard error with an average sample size of 2.95 replicates. Adapted from Walker *et al* (Walker *et al*., 2017).

Next, a clustering-based approach was employed (Figure 3.2G, see methods) to identify the timing and direction (up or downregulated) with which transcripts are *first* DE in during the experiments and categorized these as 'immediate' (within 1 h), 'fast' (1-2 h), 'downstream' (2-6 h) and 'late' (6 h+). In the pericycle, genes are almost exclusively designated as immediate or fast responders (Figure 3.2I; 48% immediate, 49% fast, 2% downstream and 1% late) although these genes may exhibit multiple changes in expression over the experiment with latter changes not captured by this approach. By contrast, patterns of expression in cortical cells are more varied with around half of transcripts not DE until more than 2 hours after the start of the experiment (Figure 3.2H; 29% immediate, 21% fast, 42% downstream and 7% late).

Whilst the majority of the transcriptome (31,524 transcripts were assayed by these arrays) does not change in expression during the timescale of these experiments in the untreated pericycle (PU) or the untreated cortex (CU) over time, thousands of transcripts do. 4,692 and 2,041 DE transcripts change over time in untreated cells of the pericycle and cortex respectively (Figure 3.2F). Amongst these DE transcripts, a high degree of cell type specificity is observed with only 291 genes being DE in both cell types. This group of 'core' transcripts includes the transcription factors *LHY* (Late elongated hypocotyl) and *CCA1* (Circadian clock-associated 1) (Figure 3.2J) which are key regulators of the plant circadian rhythm (Mizoguchi *et al*., 2002) and *PRR5* (Pseudo-Response Regulator 5), *PRR9* and *ELF4* (Early Flowering 4) which are also established to be involved with the circadian clock (Hsu and

Harmer, 2014). Also present are the histone chaperones *NRP1* (NAP1-related protein 1) and *NRP2* which are conserved throughout multicellular eukaryotic evolution. Both proteins are involved in nucleosome assembly and have roles in root proliferation (Gonzalez-Arzola *et al*., 2017; Zhu *et al*., 2006). *TORNADO*2 is also DE in both cell types and is a member of tetraspanin family of membrane proteins which are key in many aspects of plant development including regulating meristem identity and root hair patterning (Cnops *et al*., 2006). The data shows that, although the changes in gene expression in cells of the cortex and pericycle of seedlings are largely unique, both of these cell types maintain regulation over time of a number of genes that regulate essential processes including the setting of circadian rhythms and growth and developmental processes across the root.

As these cells are from untreated roots and therefore not directly responding to a stimulus, it seems that the pericycle shows more dynamic patterns of gene expression as exemplified by the greater number (2.3 times as many DE genes during the experiment) and faster nature of changes of expression (with almost all the genes DE within 2 hours) seen after dawn. This could possibly be explained by increased mitotic activity in the pericycle as the site of lateral root emergence (Casimiro *et al*., 2003) as these seedlings would be beginning to form lateral roots.

Responses to nitrogen are highly cell type specific and occur outside-to-inside: Nitrogen provision is known to stimulate root architecture changes in plants including an increase in the formation of new lateral roots (Zhang and Forde, 1998). This phenotype was replicated in these experiments (Figure 3.3A). Therefore, plants were treated with replete ammonium nitrate to identify transcripts that could be associated with root architecture remodelling. To carry out an N-treatment, seedlings were transferred to plates containing 5 mM NH_4NO_3, then roots and cells transcriptionally profiled as during the untreated timeseries before nitrogen responsive genes were computed using GP2S.

A total of 2,832 and 2,854 nitrogen-responsive DE transcripts were identified in the pericycle (PN) and cortex (CN) respectively, of which only 450 were DE in both cell types (Figure 3.3E), suggesting that the high

degree of cell type specificity seen during the baseline experiments is maintained during the response to nitrate.

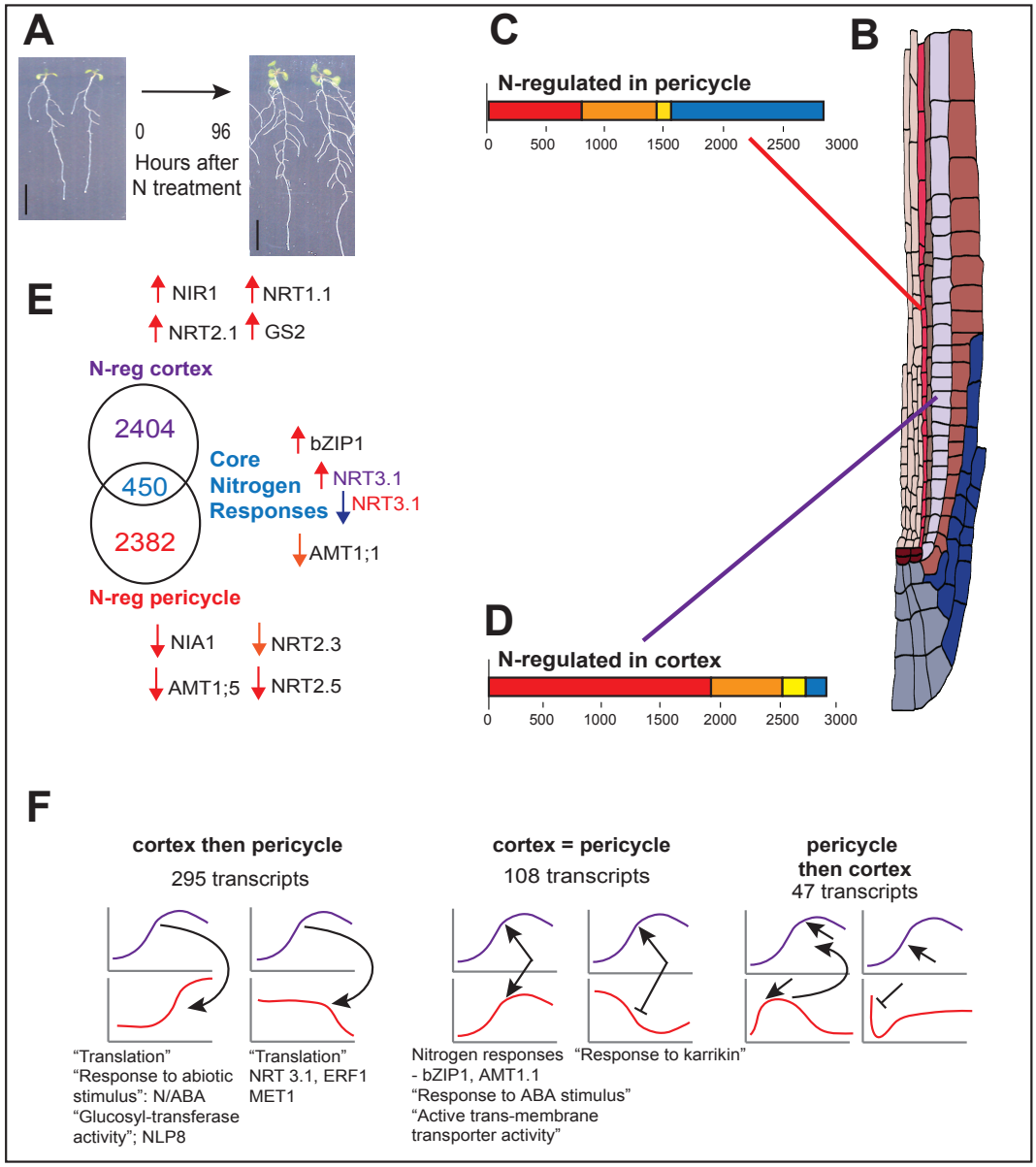

Figure 3.3. The response of pericycle and cortical cells of the *A. thaliana* root to nitrogen. (A) *A. thaliana* seedlings grown on 0.3 mM ammonium nitrate and seedlings 4 days after transfer to 5 mM ammonium nitrate. Scale bars represent 1 cm. (B-D) Number of transcripts differentially expressed in response to nitrate treatment in the pericycle (D) and cortex (C). Colours denote first statistically significant change in expression; red – immediate (0-1 h), orange – fast (1-2 h), yellow – downstream (2-6 h) and blue late (6+ h). (E) Numbers of nitrate responsive transcripts in each cell type, highlighting the 'core' of 450 transcripts that are differentially expressed in both cell types.

Markers of the response to nitrogen are annotated where the arrow indicates direction of expression (up = up-regulated, down = down-regulated) and colours are as in (C) and (D). *NRT3.1* is differentially regulated in the cortex and pericycle. (F) Delineation of response timings of the 450 nitrate-responsive genes that are DE in both cell types, highlighting a predominance of differential expression in the cortex before the pericycle. Adapted from Walker *et al* (Walker *et al*., 2017).

To validate the efficacy of the nitrogen treatment, a list of 50 commonly nitrogen-responsive genes in the root from a meta-analysis (Canales *et al*., 2014) was searched for in the expression data from PN and CN. 31 of these genes were found to be regulated in PN/CN data (7 in PN only, 13 in CN and 13 in both). Considering the data consists of only two cell types and these marker genes were identified from whole roots, this suggests that the nitrogen responses observed during PN and CN are robust but also that common nitrogen-responsive genes tend to be regulated in many root cell types.

The categorization of DE genes by clustering and their assignment to timing categories was used to help understand the response to nitrogen. Of the 2,382 nitrogen-regulated transcripts in the pericycle there is a mixture of rapid and slower responses (Figure 3.3C; 28% immediate, 22% fast, 4% downstream and 45% late) whilst the response in the cortex is dominated by early changes in gene expression (Figure 3.3D; 66% immediate, 21% fast, 7% downstream and 6% late). This suggests that the response to nitrogen is faster and more dynamic in the cortex, which is opposite to the pattern observed in untreated cells of over the same timescale (PU vs CU).

This trend for faster responses in the cortex is also apparent when looking at the ‘core’ 450 transcripts that are responsive to nitrogen in both cell types (Figure 3.3E-F). 108 of these transcripts are DE with the same timing in both cell types but not necessarily the same direction (73 show identical regulation whilst 35 are DE at the same time but in opposite directions). 295 transcripts were DE in both cell types but first in the cortex before the pericycle. 47 transcripts were DE in both cell types, but first in the pericycle. This is suggestive of a signal transduction event originating in either the epidermis or lateral root cap (which was not profiled and therefore not represented in the

timeseries data) or the cortex and being perceived within the root to at least the pericycle. Nitrate itself may be transported into the root in solution through both cell layers and through intracellular spaces since it can rapidly permeate root structures such as that of *A. thaliana* seedlings (Dechorgnat *et al*., 2011). By the first timepoint sampled after treatment (1 hour) nitrate should already be present (Krouk *et al*., 2010c) throughout the root suggesting this is a genuine signal transduction event rather than simply a delayed response due to nitrate reaching the pericycle later than the cortex.

The 450 DE 'core' nitrogen responsive genes include many known markers for the response to nitrogen (Figure 3.3E) such as *bZIP1* which is immediately upregulated in both cell types. *At*bZIP1 is transcription factor that has been demonstrated to have an important role in co-ordinating downstream responses to nitrogen (Obertello *et al*., 2010) and an inferred target of *bZIP1*, *NLP3*, is also present (Para *et al*., 2014). Also part of the core N-response is *AMT;1,1* (fast downregulated in both cell types) which functions as a high affinity ammonium transporter (Loque *et al*., 2006). A NIN-like protein, *NLP8*, is down-regulated in both cell types. *NLP8* is associated with the regulation of seed germination in response to nitrate (Yan *et al*., 2016) but mutants in this gene have also been demonstrated to exhibit reduced growth in response to nitrogen, suggesting a role in N-assimilation in seedlings (Walker *et al*., 2017).

Many genes that are associated with transport and metabolism of nitrate during the PNR are up-regulated in the cortex and/or downregulated in the pericycle. *NRT3.1*, which functions as a high affinity nitrate transporter (Okamoto *et al*., 2006) is immediately upregulated in the cortex but late downregulated in the pericycle. In the cortex, there is upregulation of the glutamine synthetase *GS2* (Cruz *et al*., 2006), the key nitrate transporters *NRT2.1* and *NRT2.2* (Li *et al*., 2007) and the nitrate reductase *NIR1* (Konishi and Yanagisawa, 2010). In the pericycle, the nitrite reductase and PNR gene *NIA1* (Konishi and Yanagisawa, 2011), is downregulated, in addition to nitrate transporters *NRT2.1* and *NRT2.5* and the ammonium transporter *AMT1;5*.

This indicates that there are differences in the response to nitrogen between the two cell types; genes associated with the PNR are typically up-regulated in the cortex but repressed in the pericycle. At the whole root level these genes are activated in response to nitrate provision (Vidal *et al*., 2015)

so the negative trend in the pericycle would be obscured if analysis was carried out at the whole organ level. This highlights the importance of cell type specific analysis for uncovering biologically relevant gene regulatory mechanisms.

Molecular changes underpinning root architecture changes in *A. thaliana* mediated by *S. meliloti*: *S. meliloti* is a N-fixing bacterium that associates with legumes such as *M. alfalfa* and *M. truncatula* and is commonly studied due to this role in nodulation. However, there is evidence that non-legumes, including *A. thaliana*, are able to perceive and respond to rhizobia and the Nod factors they produce (Liang *et al*., 2013). *S. meliloti* (rhizobia) treatment induces changes in root architecture in *A. thaliana* (Figure 3.4A-B) with inoculated seedlings producing more, but on average shorter, lateral roots, relative to uninoculated plants. Therefore, rhizobia was also included as a treatment in these experiments to ask if there were common or distinct gene expression changes underlying these two types of root architecture responses.

Following a similar experimental design to that described for the nitrogen treatment timeseries described above, the transcriptome of cortical and pericycle cells was profiled over time after inoculation with rhizobia. A total of 3,552 and 2,748 rhizobia responsive transcripts were identified in the pericycle (PR) and cortex (CR) respectively. As with the nitrate treatment, only a small (202) proportion of transcripts were regulated in both cell types (Figure 3.4C). The pericycle response exhibited only a very small proportion of changes within 2 hours (2% immediate, 2% fast, 23% downstream and 73% late) whilst the speed of changes in expression consist of a mix of early and late responses in the cortex (47% immediate, 9% fast, 10% downstream and 44% late) as was the case with the nitrate response.

Rhizobia are known to elicit short-lived transcription of genes associated with immunity in legumes during the early stages of nodulation (Kouchi *et al*., 2004; Libault *et al*., 2010; Lohar *et al*., 2006). This may be due to the presence of microbe-associated molecular patterns (MAMPs) present in the cell walls of rhizobia which could elicit a defence response despite rhizobia not being a pathogen. To test if rhizobia induced immunity in *A. thaliana*, the PR and CR data was compared to a list of differentially expressed

genes obtained from FACS-isolated cortex and pericycle cells 1 hour after treatment with 1 μM flg22 (Rich *et al*., 2019). Flg22 is a highly immunogenic peptide derived from bacterial flagellin that is able to induce expression of defence-related genes in plants (Navarro *et al*., 2004).

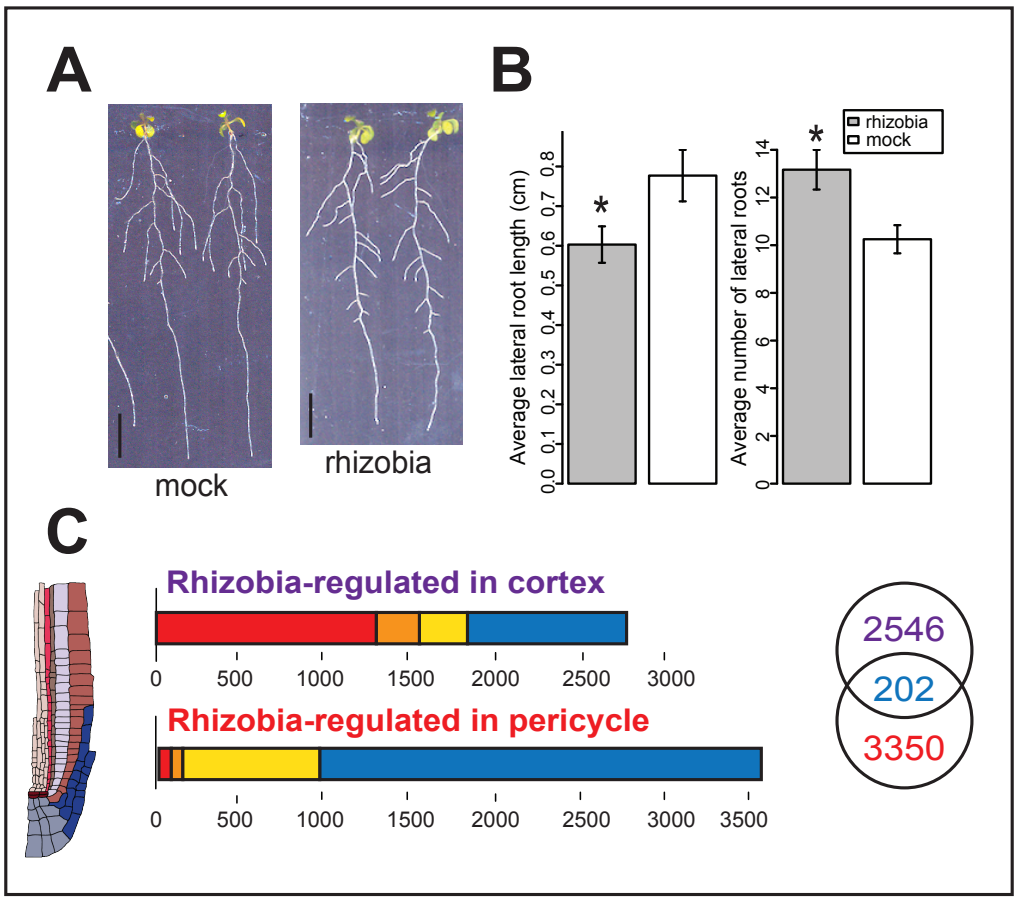

Figure 3.4. The root architecture and transcriptome response of *A. thaliana* to rhizobia. (A) *A. thaliana* seedlings 4 days after mock (left) or *S. meliloti* (right) inoculation. Scale bars represent 1 cm. (B) Average length of lateral roots (left) and average number of lateral roots (right) for *A. thaliana* seedlings 3 days after mock (white) or *S. meliloti* (grey) inoculation. n=16, error bars represent standard error and asterisks denote a significance level of P<0.05 as calculated by t-test. (C) Number of transcripts differentially expressed in response to rhizobia treatment in the cortex and pericycle. Colours denote first statistically significant change in expression; red – immediate (0-1 h), orange – fast (1-2 h), yellow – downstream (2-6 h) and blue late (6+ h). Adapted from Walker *et al* (Walker *et al*., 2017).

The Rich *et al.* 2019 dataset (biorx number 302448) consisted of a list of 166 flg22-responsive genes in the cortex and 121 flg22-responsive in the pericycle. Only a small proportion of these were also found to be DE in PR/CR data (pericycle – 8 of 121; 7% and cortex - 38 of 166; 23%). However, some marker genes of defence responses are present in both the PR and CR response lists, including FLS2. FLS2 is the primary receptor for flg22 and is responsible for inducing innate immunity downstream of flg22 perception (Sun *et al*., 2013). Interestingly, the flagellin of another rhizobial species, M. loti, has been demonstrated not to possess MAMP activity in the flg22-sensitive legume L. japonicus (Lopez-Gomez *et al*., 2012) but it is unclear if this is the case for the flagellin of other rhizobia. Of the genes common between both flg22- and rhizobia-response datasets in the cortex, the majority (27 of 38; 71%) were determined to be immediately rhizobia responsive on the basis of their cluster membership, in line with rapid activation in response to flg22. Of the genes conserved in both datasets in the pericycle, all but one are classified as 'late'. Collectively, the response to rhizobia in the cortex appears to partially resemble a plant defence response whilst the pericycle response to rhizobia is distinct from the flg22 response.

Expression of cell identity markers is maintained during environmental responses: Based on the numbers of DE transcripts, these experiments suggest that although there are common elements of environmental responses between cell types, cell types maintain their identities following a change in the environment. To further explore this, all lists of DE genes were queried with lists of genes that are defined as 'cell type specifically enriched' (CTSE) genes for the cortex and the pericycle (Bargmann *et al*., 2013b). CTSEs are genes that exhibit preferential expression in a single cell type and therefore can be considered as an extension of cell type identity. This dataset included 85 and 168 genes identified as CTSEs in the pericycle and cortex respectively. However, these lists were constructed at the gene level, rather than the gene model level as in the data presented here. Therefore, the Bargmann lists were expanded (and three genes not present on the microarray were removed; At1g24280, At4g36670 and At5g42600) to include all gene models of these CTSEs represented on the microarrays. This gave a

total of 104 CTSE transcripts in the pericycle and 194 in the cortex to use for analysis.

A significant proportion of the CTSEs from both cell types (Figure 3.5A-B) were not differentially expressed during any of the timeseries experiments, and also retained their cell type specific expression – their expression was enriched in the cell type that they were identified in in the Bargmann *et al.* 2013 work. However, many (60/104 in the pericycle; 58% and 110/194 in the cortex; 57%) did show differential expression during at least one experiment, including a small proportion (pericycle 7%, cortex 14%) of CTSEs identified as cortex- or pericycle-associated that were nitrogen/rhizobia responsive in the other cell type. This does not necessarily suggest that cell identity is affected during the treatments, since these genes still predominantly showed preferential cell type specific expression. However, it does suggest that CTSEs (and therefore cell identity) could exhibit plasticity during environmental changes.

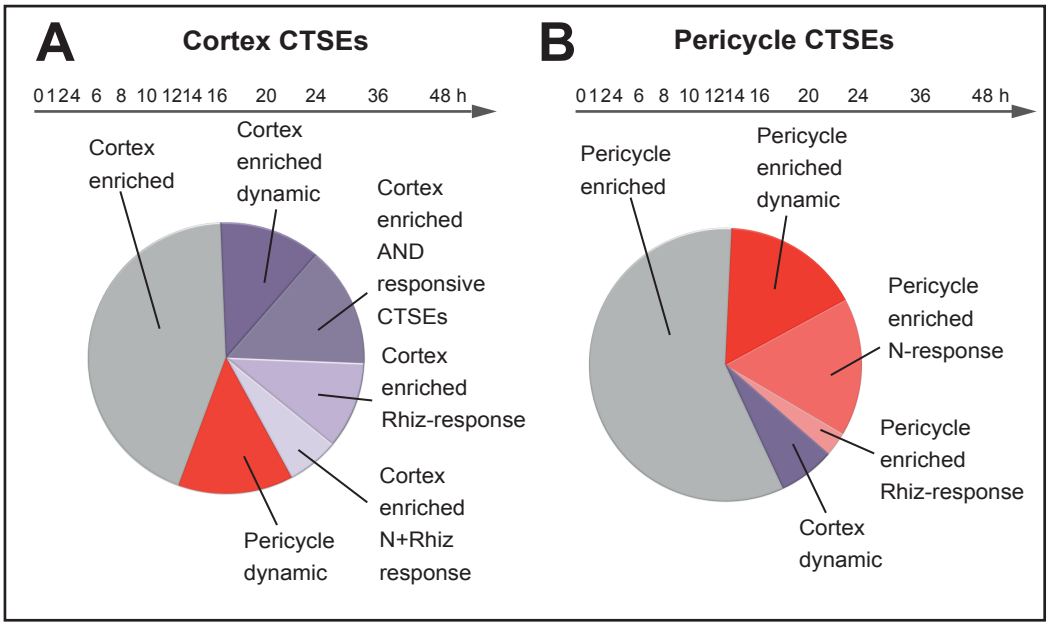

Figure 3.5: Expression of cortex and pericycle cell type specific enriched genes during environmental change. A-B) Charts showing the proportion of 194 transcripts identified as cortex (A) CTSEs and 100 transcripts identified as pericycle CTSEs (B). Transcripts annotated as 'enriched' do not show differential expression within any time series whilst those annotated as 'dynamic' to a cell type show differential expression in the untreated timecourse for that cell type. Transcripts labelled as 'responsive' are DE

following nitrogen or rhizobia treatment. Adapted from Walker *et al* (Walker *et al*., 2017).

Overall, this reinforces the hypothesis that the cortex and pericycle both maintain their distinct cell identities but these are able to be modulated during environmental responses.

Understanding how cortical and pericycle gene expression is co-ordinated using network inference: An interpretation of the data presented above is that the pericycle and the cortex show a very high degree of cell type specificity in terms of their gene expression in untreated cells and that this is maintained during responses to nitrogen and rhizobia. But neither cell type is directly exposed to either stimulus and it therefore seems unlikely that these highly cell type specific responses could occur without some degree of co-ordination at the whole root level. To identify regulators of these responses, causal transcription networks were generated (see methods) for all the experiments. Causal transcription networks can be inferred from timeseries data to predict regulatory interaction between differentially expressed genes and thus help in visualizing how responses are regulated. Since the network inference algorithm has to have samples with consistent spacing between timepoints (sampling every 2 hours), only data up to 16 hours after inoculation were used for network inference. Previously described analysis of the timing of gene responses (where most responses at least first occurred in the first few hours) suggests that this time period contains the most pertinent information for understanding these responses.

Network inference assigns transcripts as regulatory nodes if they are transcriptional regulators. In total 2,460 genes (2,946 transcripts) from the arrays used were assigned to be putative regulators. This was based on known or putative TF activity using data from the NCBI Conserved Domains database indicating the presence of conserved protein domains indicative of DNA binding combined with gene annotations (Gene Ontology and TAIR), indicative of regulatory effect.

Each network connected hundreds of genes via regulatory connections, or ‘edges’ (Table 3.1). The networks inferred from gene expression data in the

cortex show different network architectures to those for the pericycle, based on analysis of the network degree, betweenness and connectivity. In spite of the networks in the pericycle and cortex consisting of similar edge to node ratios, the pericycle networks tend to be dominated by fewer, larger nodes although this would be expected to some extent given the higher number of nodes in some of the pericycle networks. This is particularly evident when looking at the largest hubs in each network. A hub can be defined as a node with a large number of inferred targets. The three largest hubs in each pericycle network (PU; 71, 66 and 53, PN; 76, 75 and 48, PR 84, 44 and 39) are larger than all of the equivalent 3 hubs in the cortex networks (CU; 37, 36 and 34 CN; 34, 27 and 24, CR 28, 24 and 17).

Network	PU	PN	PR	CU	CN	CR
Nodes	1651	1557	785	898	981	892
Edges	1974	2018	1099	1271	1348	1147
DAP-seq validation	61 of 315	52 of 246	48 of 216	64 of 235	73 of 100	61 of 289
Treatment responsive	N/A	670 (43%)	388 (49%)	N/A	770 (78%)	677 (76%)
Response timing (%I/F/D/L)	N/A	1/2/5/82	17/26/5/51	N/A	66/18/7/9	55/9/5/31

Table 3.1: Statistical analysis of network inference. Comparison of features of networks inferred from each timecourse experiment across a range of categories. **Nodes:** number of regulators present in each network. **Edges:** number of regulatory actions inferred between nodes in each network. **DAP-seq validation:** Number of edges validated in each network by a DAP-seq dataset (O'Malley *et al*., 2016) expressed as a proportion of the total number of edges present in both the O'Malley dataset and the network for which the comparison is being made. **Treatment responsive:** the proportion of nodes in each network that were identified as exhibiting significant DE relative to untreated cells. **Response timing:** percentages of nodes that were designated immediate (0-1 h), fast (1-2 h), downstream (2-6 h) or late (6+ h)

responders to treatment expressed as a proportion of total treatment responsive genes in that network.

To validate the node-node edge connections inferred from expression data, edges predicted in the networks were compared with those identified from a study predicting transcription factor binding sites using DAP-seq (O'Malley *et al*., 2016). Although many edges in the networks were not tested in this study, a proportion of edges were present in both datasets (Table 3.1). Across the networks, between 19-73% of interactions were conserved giving confidence in the edge prediction by network inference when the absence of whole root data from these experiments is considered.

The networks were then analysed to understand how processes change within the cortex and pericycle over time and in response to the nitrogen and rhizobial treatments. The PU (Figure 3.6) and CU (Figure 3.7) networks are constructed from expression data of cells not responding to a specific treatment, and therefore these networks provide an insight into developmental changes occurring in the cortex and pericycle of growing seedlings. The key plant circadian clock genes *CCA1* and *LHY* are represented in both networks. *CCA1* has 3 targets in the PU network and 11 in the CU network whilst *LHY* has 12 and 7 respectively. The genes inferred to be regulated by both *CCA1* and *LHY* are unique in both cell types, highlighting that both cell types have distinct identities.

The nitrate-responsive gene *TGA1* was found to be the fifth-largest hub in the CU network with 32 outgoing edges and 2 incoming edges. The two TFs inferred to regulate *TGA1* are *WRKY36* (positively) and *TGA4* (negatively). This is curious because *TGA1* and *TGA4* are associated with the PNR (Alvarez *et al*., 2014) and defence responses (Kesarwani *et al*., 2007) although this does not exclude other roles. Both proteins are thought to be functionally redundant so the idea that *TGA4* is a transcriptional repressor of *TGA1* is novel and may represent a negative feedback mechanism. Another gene associated with nitrate responses, *HHO1* (1 regulator, 3 targets), is present in the CU network but not the PU network, suggesting that it has a cell type specific role.

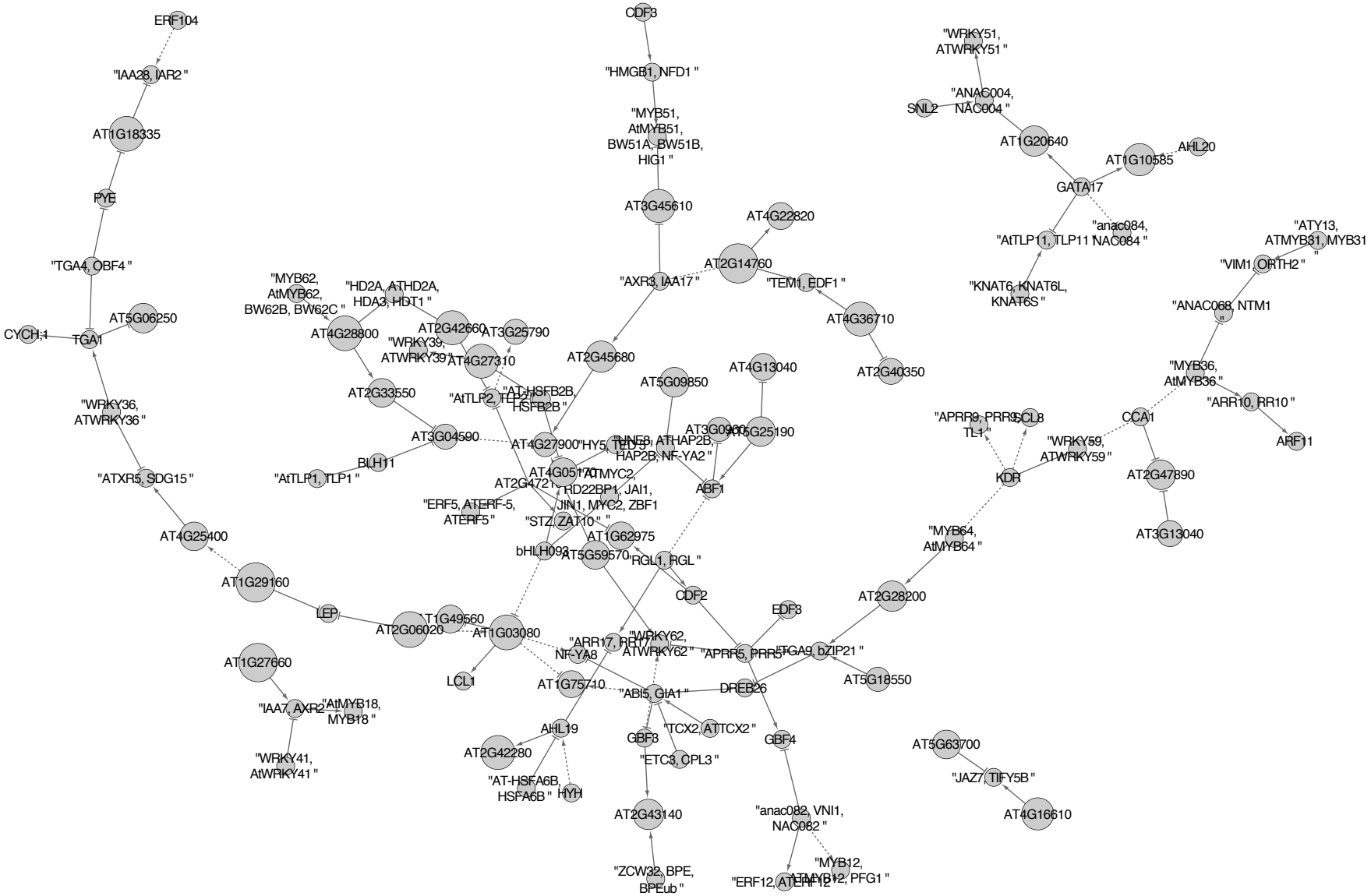
ERF104
"IAA28, IAR2"
AT1G18335
PYE
"TGA4, OBF4"
CYCH;1
TGA1
AT5G06250
"WRKY36, ATWRKY36"
"ATXR5, SDG15"
AT4G25400
AT1G29160
LEP
AT2G06020
T1G49560
AT1G27660
"IAA7, AXR2"
"AtMYB18, MYB18"
"WRKY41, AtWRKY41"
ODF3
"HMGB1, NFD1"
"MYB51, AtMYB51, BW51A, BW51B, HIG1"
AT3G45610
"AXR3, IAA17"
AT4G22820
AT2G14760
"TEM1, EDF1"
AT4G36710
AT2G40350
"MYB62, AtMYB62, BW62B, BW62C"
"HD2A, ATHD2A, HDA3, HDT1"
AT4G28800
AT2G42660
"WRKY39, ATWRKY39"
AT3G25790
AT4G27310
AT2G45680
AT2G33550
"AtTLP2, TLP2"
"AT-HSFB2B, HSFB2B"
AT3G04590
BLH11
"AtTLP1, TLP1"
AT4G27900
"HY5, TED3"
TINF8, ATHAP2B, HAP2B, NF-YA2
AT4G05170
ATMYC2, RD22BP1, JAI1, JIN1, MYC2, ZBF1
AT2G47210
"ERF5, ATERF-5, ATERF5"
"STZ, ZAT10"
AT5G09850
AT4G13040
AT3G09600
AT5G25190
ABF1
AT1G62975
bHLH092
AT5G59570
"RGL1, RGL"
CDF2
"WRKY62, ATWRKY62"
"ARR17, RER7"
NF-YA8
LCL1
AT1G75710
"ABI5, GIA1"
DREB26
EDF3
"APRR5, PRR5"
"TGA9, bZIP21"
AT5G18550
AHL19
AT2G42280
"AT-HSFA6B, HSFA6B"
HYH
GBF3
"TCX2, ATTCX2"
"ETC3, CPL3"
GBF4
AT2G43140
"ZCW32, BPE, BPEub"
"anac082, VNI1, NAC082"
"MYB12, ATMYB12, PFG1"
"ERF12, ATERF12"
"WRKY51, ATWRKY51"
"ANAC004, NAC004"
SNL2
AT1G20640
AT1G10585
AHL20
GATA17
"anac084, NAC084"
"AtTLP11, TLP11"
"KNAT6, KNAT6L, KNAT6S"
"ATY13, ATMYB31, MYB31"
"VIM1, ORTH2"
"ANAC068, NTM1"
"MYB36, AtMYB36"
"ARR10, RR10"
ARF11
"APRR9, PRR9, TL1"
SCL8
CCA1
"WRKY59, ATWRKY59"
KDR
AT2G47890
AT3G13040
"MYB64, AtMYB64"
AT2G28200
AT5G63700
"JAZ7, TIFY5B"
AT4G16610

Figure 3.6: Causal transcription network of changes over time in untreated pericycle cells. For simplicity, the network is visualized as a regulatory backbone with only nodes displayed that are predicted to regulate at least one target and with the size of nodes increasing with the number of outgoing edges. Nodes not connected to the main network with fewer than 2 regulatory interactions are omitted. Edges with arrow heads denote positive regulation and edges with flat heads denote negative regulation.

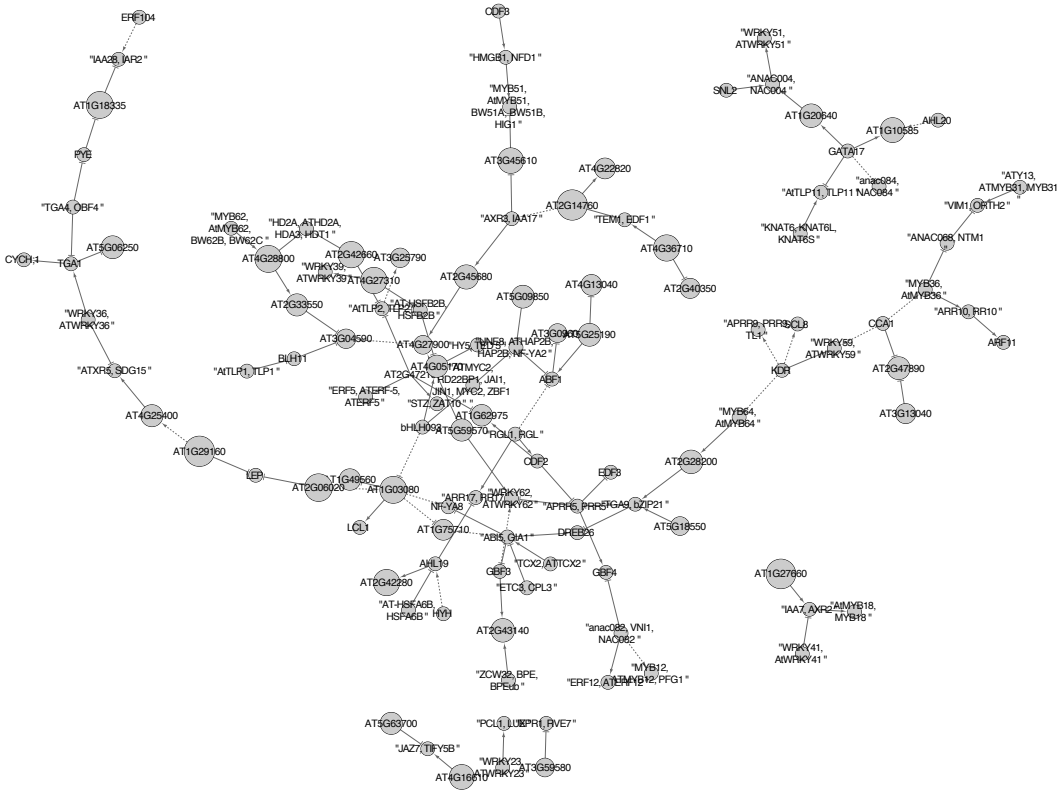

Figure 3.7: Causal transcription network of changes over time in untreated cortex cells. For simplicity, the network is visualized as a regulatory backbone with only nodes displayed that are predicted to regulate at least one target and with the size of nodes increasing with the number of outgoing edges. Nodes not connected to the main network with fewer than 2 regulatory interactions are omitted. Edges with arrow heads denote positive regulation and edges with flat heads denote negative regulation.

Insights into co-ordination of nitrate responses from PN/CN networks: Networks for nitrate and rhizobia responses were constructed based on genes DE within a time series (DE genes calculated by BATS) as opposed those DE between time series (those calculated by GP2S). This was because network connections over time within each treatment could involve genes that change in response to treatments but also genes that must be present but whose expression does not significantly change after treatment. Therefore, to enable visualization of the genes that do change due to treatment, these nodes were colour-coded in the networks to distinguish between this. GOrilla (Eden *et al*., 2009) was then used to query the networks for overrepresentation of GO terms. To focus on understanding the higher order of regulation of the rhizobial and nitrogen responses, networks were visualized as 'regulatory backbone' networks where only nodes with at least one target are displayed and the size of the shown nodes scales with larger nodes predicted to have more targets.

Identification of nitrogen or rhizobia responsive genes within their respective networks reveals a much higher proportion of genes regulated in the cortex networks to be treatment responsive than in the pericycle (Table 3.1). This suggests that environmental responses in the pericycle potentially show less plasticity than those in the cortex. The N-response in the pericycle (Figure 3.6B) seems to be regulated by many genes that are not directly N-responsive but are nonetheless are important in co-ordinating the response. The majority of genes in the cortex network (Figure 3.6A) are N-responsive. When analysing response timings, a much higher proportion of expression changes occur first within 2 hours of treatment (CN: 84% vs PN: 3% and CR: 64% vs PR: 43%) in the cortex than the pericycle (Table 3.1). This shows that the trend for faster responses to nitrogen or rhizobia seen amongst DE genes of the cortex relative to the pericycle (Figure 3.3B, Figure 3.3D, Figure 3.4C) is also conserved in their respective networks.

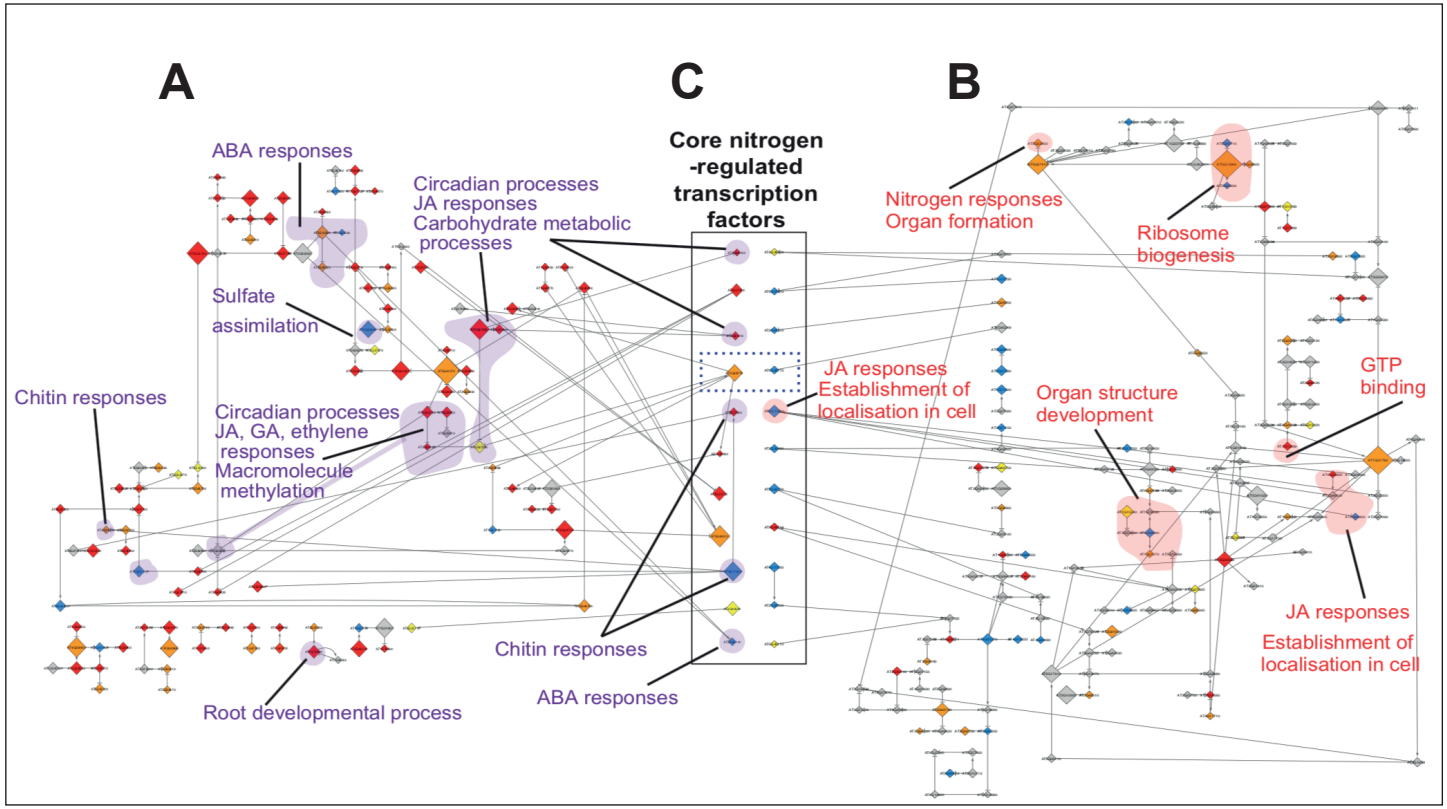

Figure 3.8: Co-ordination of the nitrate response in the cortex and pericycle. A-B) Regulatory backbone visualization of inferred casual transcription networks in the cortex (A) and pericycle (B) during response to nitrate. The network is constrained to only show nodes that have at least one outgoing edge and the size of nodes is scaled with larger nodes possessing more outgoing edges. C) Conserved transcription factor hubs present in both networks are arranged together. Overrepresentation of gene ontologies amongst a hub and its targets is denoted by shading. Colours denote first statistically significant change in expression; red – immediate (0-1 h), orange – fast (1-2 h), yellow – downstream (2-6 h) and blue late (6+ h). Edges with arrow heads denote positive regulation and edges with flat heads denote negative regulation. Adapted from Walker *et al* (Walker *et al*., 2017).

Because genes involved in the PNR are well characterized in *A. thaliana*, it is possible to query the PN/CN networks for these genes. This means that the PNR, which is a process co-ordinated across the whole root can be assessed at the cell type level. bZIP1 is present and functions as a key hub in both cortex and pericycle networks (7 outgoing edges in PN and 17 in CN). Most of the targets of bZIP1 in either network are not functionally associated with nitrate responses but the ammonium transporter AMT1;2 is repressed in the cortex, which is consistent with its downregulation of genes involved in ammonium transport seen elsewhere in the data. The NIN-like protein *NLP3* is present in both networks (2 targets in PN, 17 in CN). NIN-like proteins are a family of TFs thought to be evolutionarily related to highly conserved legume symbiosis regulator *NIN* and are known to be important in nitrate responses in *A. thaliana* (Konishi and Yanagisawa, 2013).

Many genes associated with PNR are present in only one network, suggesting that their effects are actually cell type specific. *NRT3.1* is up-regulated only in the CN network by two TFs, one of which is another NIN-like protein, *NLP9*. However, *NLP9* itself is also present in the pericycle where it has 30 inferred targets, suggesting it is a key hub gene in the nitrogen response in this cell type, despite not being differentially expressed between the PU and PN timeseries. Also exclusive to the PN network is *TGA1* which has 11 targets. The presence of *TGA1* in the PN network is noteworthy because it is also one of largest hubs in CU network but is not present in the PU network at all, suggesting it may have very different roles between the two cell types, with a role in the N-response in the pericycle but not the cortex. Also, of particular interest is the expression of *HRS1* and its homolog *HHO1*. *HRS1* is only present in the PN network where it is induced by one TF and repressed by another and has six targets of its own. *HHO1* however is only connected in the CN network where it has 8 outgoing edges.

Not present in either network are numerous genes that which are known to be markers of the PNR including *NRT1.1*, *NRT2.1*, *NRT2.2*, *NIA1*, *NIA2*, and *NIR1* (Vidal *et al*., 2015). It is likely that these genes are missing since they are part of whole root responses that are more strongly expressed in other cell types such as the epidermis or endodermis which are absent in this study.

Because these networks were formed from genes DE within an experiment, the clock genes *CCA1* and *LHY* are present in the nitrate networks of both cell types. Using the combination of BATS and GP2S DE lists it is possible to discriminate between circadian influences on rhizobial and N responses and changes in gene expression due to circadian rhythms in general. *LHY* only has 2 targets in the PN network and both are not nitrogen-responsive but 17 of 21 targets in the CN network are. For *CCA1*, 1 of its 6 targets is nitrogen responsive in the PN network whilst 8 of 9 targets are N-regulated in the CN network. This would suggest that these two clock genes assume additional regulatory roles during nitrogen responses in the cortex.

When evaluating the PN/CN networks to understand the plant's response to N (Figure 3.8A-B) only a small proportion (222) of the total nodes (981 and 1557 in the cortex and pericycle respectively) in each network are conserved between both networks. However, when nodes and their targets are queried for overrepresented GO terms, in order to assess to what extent biological processes and molecular functions are distinctly regulated (Figure 3.8). Some similar processes are observed in both cell types; for instance, there are two different modules in both networks which regulate distinct sets of genes associated with JA responses. There are 11 transcription factors are conserved between, and are N-responsive in, the CN and PN networks (Figure 3.8C). The majority (8 of 11) of these showed a cortex then pericycle preference in terms of timing of expression changes, supporting the proposed mechanism discussed earlier of signal transduction through the cell types of the root during the response to nitrate (Figure 3.3F).

This suggests that different sets of genes, which may be considered functionally redundant at the whole root level, regulate common functions within the two cell types and therefore can be considered functionally distinct on the basis of the timing and spatial location of their expression and this assertion is supported by data at the gene family level (Walker *et al*., 2017).

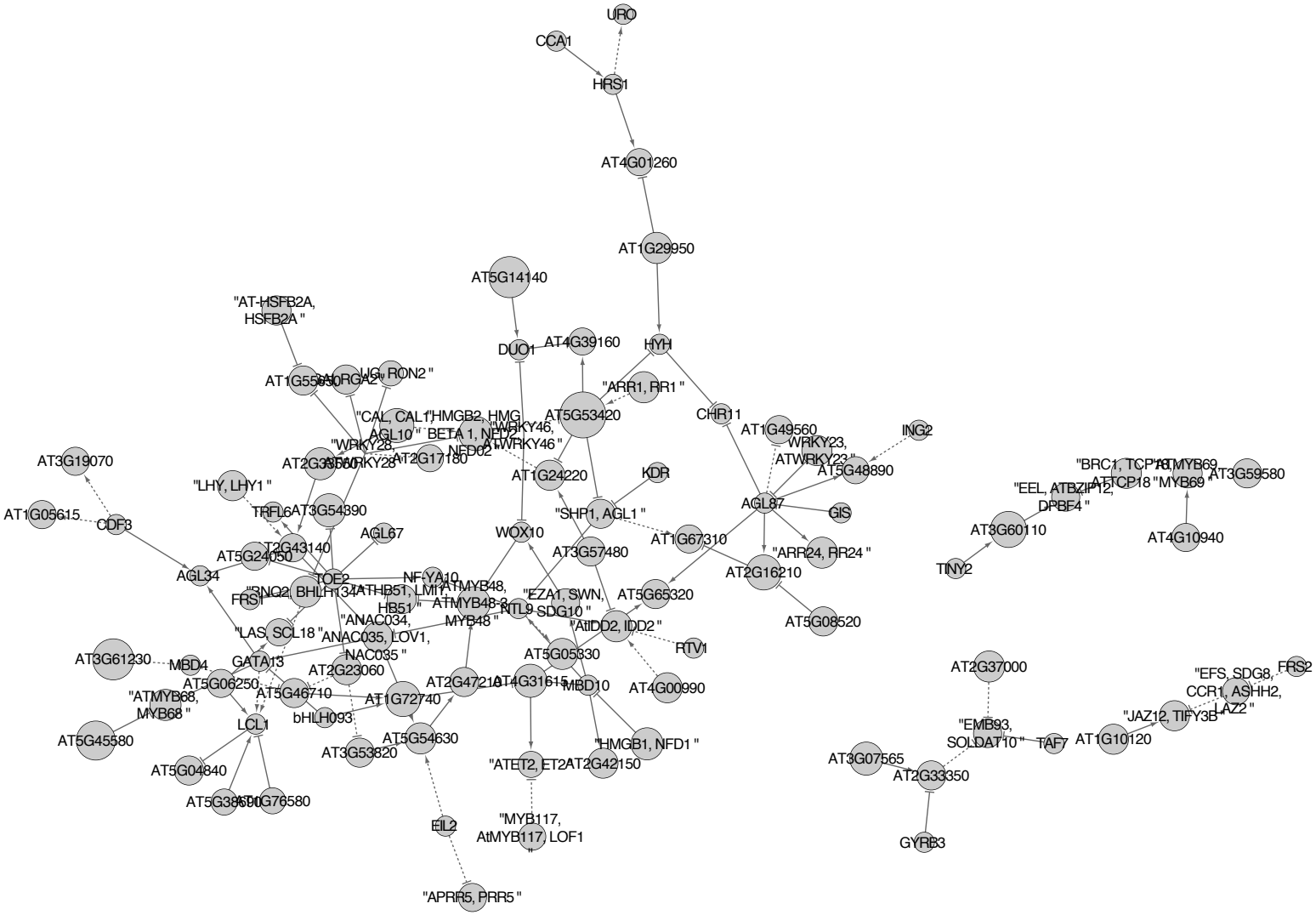
URC0
CCA1
HRS1
AT4G01260
AT1G29950
AT5G14140
"AT-HSFB2A, HSFB2A "
DUO1
AT4G39160
HYH
"ARR1, RR1 "
AT5G53420
CHR11
AT1G49560
ING2
WRKY23, ATWRKY23 "
AT5G48890
AT3G19070
"LHY, LHY1 "
AT1G05615
CDF3
TRFL6
AT3G54390
AT2G43140
AT5G24050
AGL34
AGL67
AT1G24220
KDR
"SHP1, AGL1 "
AGL87
GIS
WOX10
AT3G57480
AT1G67310
"ARR24, RR24 "
AT2G16210
AT5G08520
"BRC1, TCP18 "
ATMYB69
AT3G59580
ATTCP18
"MYB69 "
"EEL, ATBZIP12, DPBF4 "
AT3G60110
AT4G10940
TINY2
TOE2
NF-YA10
ATMYB48, ATHB51, LMI1, HB51 "
ATMYB48-2
"EZA1, SWN, SDG10 "
RTL9
AT5G65320
"AtIDD2, IDD2 "
RTV1
"ANAC034, MYB48 "
"ANAC035, LOV1, NAC035 "
"LAS, SCL18 "
FRS1
"BNQ2, BHLH134 "
AT3G61230
MBD4
GATA13
AT5G06250
AT5G46710
AT2G23060
AT1G72740
AT2G47210
AT4G31615
MBD10
AT4G00990
AT5G05330
"ATMYB68, MYB68 "
AT5G45580
LCL1
bHLH093
AT3G53820
AT5G54630
"HMGB1, NFD1 "
"ATET2, ET2"
AT2G42150
AT5G04840
AT5G38600
AT5G76580
EIL2
"MYB117, AtMYB117 LOF1 "
"APRR5, PRR5 "
AT2G37000
"EFS, SDG8, CCR1, ASHH2, LAZ2 "
FRS2
"JAZ12, TIFY3B "
"EMB93, SOLDAT10 "
TAF7
AT1G10120
AT3G07565
AT2G33350
GYRB3

Figure 3.9: Causal transcription network of changes over time in pericycle cells treated with *S. meliloti*. For simplicity, the network is visualized as a regulatory backbone with only nodes displayed that are predicted to regulate at least one target and with the size of nodes increasing with the number of outgoing edges. Nodes not connected to the main network with fewer than 2 regulatory interactions are omitted. Edges with arrow heads denote positive regulation and edges with flat heads denote negative regulation.

Analysis of networks of rhizobia-responsive transcripts: The response to rhizobia in *A. thaliana* cortical cells seems to share some elements with defence responses as determined by comparison with flg22-regulated gene expression. Because this cannot be said to be true of the pericycle, the PR (Figure 3.9) and CR (Figure 3.10) networks were examined more closely.

In line with finding both clock genes *CCA1* and *LHY* to be DE in both CR and PR timeseries, they were both found to be within both PR and CR networks. *CCA1* has 10 targets in PR (1 of which is also rhizobia-responsive) and 7 in PR (6 of which are also rhizobia-responsive) whilst *LHY* has 5 (1 of which is also rhizobia-responsive) and 9 (8 of which is also rhizobia-responsive) respectively. This supports the findings of analysis of the N-response networks in these transcription factors are involved in environmental responses in the cortex but not in the pericycle whilst also reflecting the wider trend that a higher proportion of genes in the CN/CR network are treatment-responsive than their counterparts in the pericycle (Table 3.1).

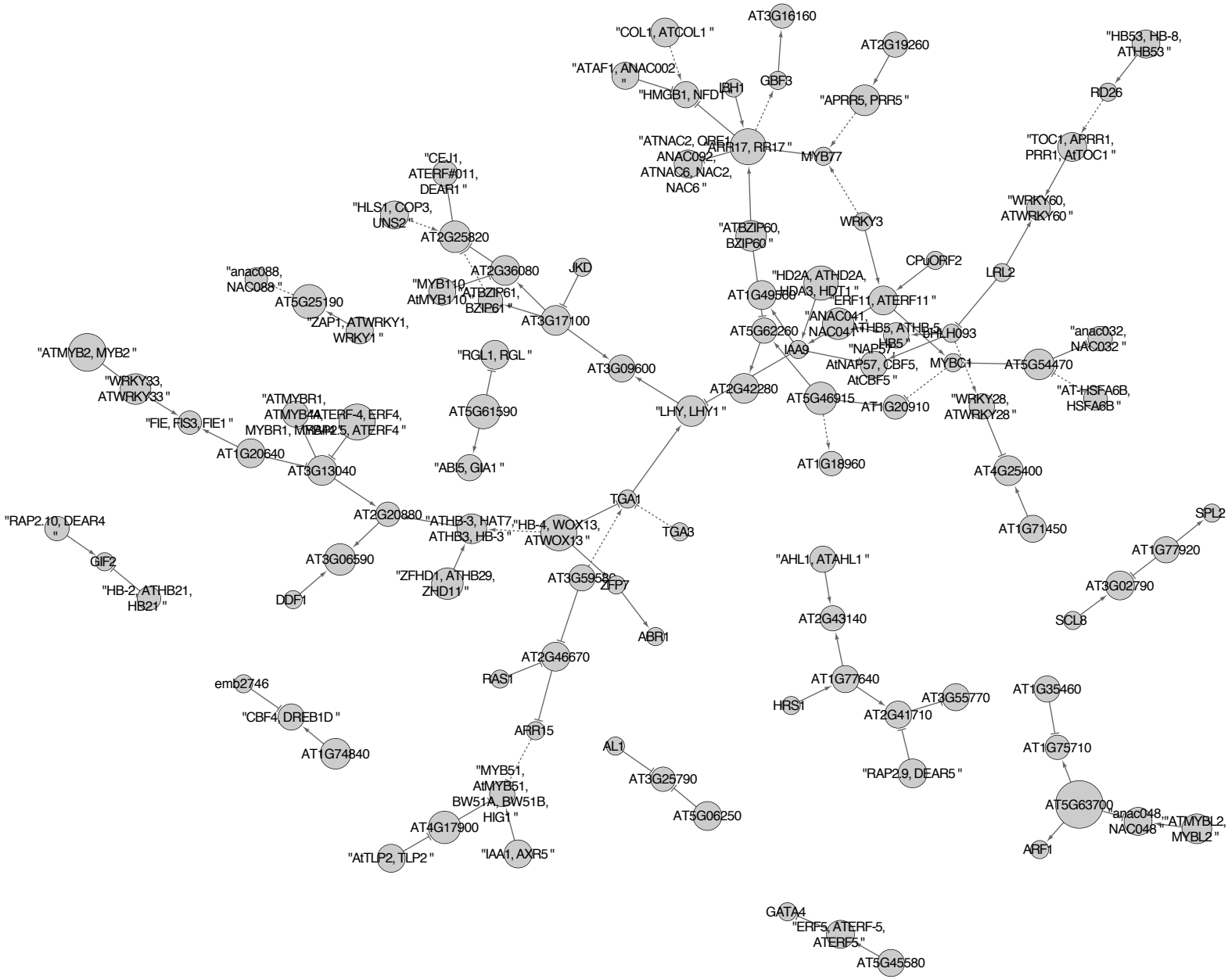
AT3G16160
"COL1, ATCOL1"
AT2G19260
"HB53, HB-8, ATHB53"
"ATAF1, ANAC002"
GBF3
IBH1
"HMGB1, NFD1"
"APRR5, PRR5"
RD26
"ATNAC2, ORE1, ANAC092, ATNAC6, NAC2, NAC6"
ARR17, RR17"
MYB77
"TOC1, ARR1, PRR1, AtTOC1"
"CEJ1, ATERF#011, DEAR1"
"HLS1, COP3, UNS2"
"ATBZIP60, BZIP60"
WRKY3
"WRKY60, ATWRKY60"
AT2G25820
CPuORF2
LRL2
AT2G36080
JKD
"anac088, NAC088"
"MYB110, AtMYB110"
"ATBZIP61, BZIP61"
AT5G25190
"ZAP1, ATWRKY1, WRKY1"
AT3G17100
AT1G49560
"HD2A, ATHD2A, HDA3, HDT1"
"ERF11, ATERF11"
"ANAC041, NAC041"
AT5G62260
ATHB5, ATHB-5, HB5
bHLH093
"anac032, NAC032"
"ATMYB2, MYB2"
"RGL1, RGL"
"NAP57"
IAA9
AT3G09600
MYBC1
AT5G54470
"WRKY33, ATWRKY33"
AtNAP57, CBF5, AtCBF5"
"AT-HSFA6B, HSFA6B"
"ATMYBR1, ATMYB4A, MYBR1, MYB4"
"ATERF-4, ERF4, MRAP2.5, ATERF4"
AT5G61590
AT2G42280
AT5G46915
AT1G20910
"WRKY28, ATWRKY28"
"FIE, FIS3, FIE1"
"LHY, LHY1"
AT1G20640
AT3G13040
"ABI5, GIA1"
AT1G18960
AT4G25400
TGA1
AT2G20880
"ATHB-3, HAT7, ATHB3, HB-3"
"HB-4, WOX13, ATWOX13"
TGA3
AT1G71450
SPL2
"RAP2.10, DEAR4"
AT3G06590
AT1G77920
GIF2
"ZFHD1, ATHB29, ZHD11"
AT3G59580
ZFP7
"AHL1, ATAHL1"
AT3G02790
"HB-2, ATHB21, HB21"
DDF1
ABR1
AT2G43140
SCL8
AT2G46670
RAS1
AT1G77640
AT3G55770
AT1G35460
emb2746
HRS1
AT2G41710
"CBF4, DREB1D"
ARR15
AT1G75710
AT1G74840
AL1
"RAP2.9, DEAR5"
"MYB51, AtMYB51, BW51A, BW51B, HIG1"
AT3G25790
AT5G63700
"anac048, NAC048"
"ATMYBL2, MYBL2"
AT4G17900
AT5G06250
ARF1
"AtTLP2, TLP2"
"IAA1, AXR5"
GATA4
"ERF5, ATERF-5, ATERF5"
AT5G45580

Figure 3.10: Causal transcription network of changes over time in cortex cells treated with *S. meliloti*. For simplicity, the network is visualized as a regulatory backbone with only nodes displayed that are predicted to regulate at least one target and with the size of nodes increasing with the number of outgoing edges. Nodes not connected to the main network with fewer than 2 regulatory interactions are omitted. Edges with arrow heads denote positive regulation and edges with flat heads denote negative regulation.

The PNR-associated *HRS1* transcription factor is present in both CR and PR networks. In the pericycle *HRS1* is inferred to be transcriptionally induced by *CCA1* and has 44 predicted targets. In the cortex, *HRS1* does not have any regulators and 4 targets. *HRS1* homolog *HHO1* is present in the CR network (2 regulators, 2 targets) meaning it is connected in all three cortical networks but absent in all pericycle ones. Both of these genes have previously been assumed to exhibit functional redundancy (Medici *et al*., 2015) but this suggests that their activities may be distinct on the basis of their expression in different cell types.

Conclusions: FACS was used to isolate cells of the cortex and the pericycle at a range of timepoints up to 48 hours after treatment with replete nitrate or rhizobia, both of which are able to influence plant root architecture. Transcriptomics of these cells over time was used to identify changes in gene expression that allowed examination of molecular changes that might underpin root architecture reshaping.

Based on the numbers of DE transcripts, both cell types have very strong cell type response identities as there is very little overlap between the two cell type responses. It was also seen that large differences exist in DE gene activity between treatments for each cell type. Together, this suggests that the identity of a cell type is flexible and is able to be modulated by the environment. This is reinforced by the findings from analysis of the expression of putative cell type specific enriched (CTSE) genes (Bargmann *et al*., 2013b) that they are cell-type enriched but also responsive to environmental change.

Analysis of the data also suggest the possibility of signal transduction within the root during environmental responses which originates in outer layers

of the root and travels inwards (cortex to pericycle in this case). The inclusion of analysis of responses of more cell types, particularly the epidermis (the outer most layer of the root) and the endodermis (which lies in between the cortex and the pericycle) would facilitate further study of this possible signal transduction. Key transcriptional regulatory hubs of pericycle and cortex cells were identified from the network analysis in untreated conditions and during the response to nitrate and rhizobia. Some genes assumed to be functionally redundant, such as *HRS1*/*HHO1* seem to be preferentially expressed in one cell type. Therefore, these genes which might appear to be functionally redundant may actually be functionally 'separate' on the basis of their patterns of expression. This could be resolved by cell type-specific analysis of mutants for genes of interest (for instance, *hrs1*/*hho1* mutants) to see how gene regulatory networks are perturbed.

This approach, using FACS to generate longitudinal and cell type specific gene expression data, has uncovered trends during root development which are obscured at the whole root level. This highlights the importance of analyzing the expression of genes through both time and space in order to fully understand how the timing and localization of this might moderate their activity as previous studies of root development (Birnbaum *et al*., 2003; Brady *et al*., 2007) and environmental responses have found (Geng *et al*., 2013; Iyer-Pascuzzi *et al*., 2011).

Chapter 4: Molecular changes underpinning 'friend vs. foe' recognition in legume roots

Introduction:

Complex interactions between the legume nodulation machinery and the plant immune system are required for establishment of root nodules: The rhizosphere is an extremely complex environment in which plant roots are continuously exposed to a diverse range of microorganisms (Berendsen *et al*., 2012). Whilst the majority of these are harmless to a plant, pathogens may also be present amongst this population. Plants possess a basal innate immune system as their primary defence against pathogens. This first layer of immunity, called MTI, consists of cell surface receptor kinases (PRRs) that recognize widely conserved chemical features associated with bacteria, fungi and other possible pathogens. Stimulation of these receptors activates their kinase activity and begins an intracellular signaling cascade that results in an immune response (Boller and Felix, 2009).

As the molecular patterns recognized by these receptors are often conserved amongst pathogens and non-pathogens, innocuous or even symbiotic bacteria can also induce immune responses. In the case of plant-microbe symbiotic interactions such as nodulation, this can potentially inhibit formation of mutualistic interactions. A wide range of immune elicitors and defence hormones have been shown to inhibit nodulation (Lopez-Gomez *et al*., 2012; Stacey *et al*., 2006; Sun *et al*., 2006). A sustained immune response to rhizobia, preventing rhizobial colonization, would result in the benefits of the interaction to both parties being lost. Consequentially, legumes have had to evolve a system allowing the selective lack of development of immune responses in response to *bona fide* symbionts without providing a potential route to be exploited by pathogens (Berrabah *et al*., 2015; Cao *et al*., 2017).

There is evidence to suggest that rhizobia and signaling molecules associated with them actually do exert a short-lived activation of legume defences during the early stages of nodulation, although the density of inoculum in these studies is typically greater than would be encountered outside of laboratory conditions. Transient up-regulation of genes associated

(based on ontology) with defence responses shortly after inoculation with rhizobia or Nod factors has been demonstrated across several model legume species including *L. japonicus*, soybean (*Glycine max*) and *M. truncatula* (Kouchi *et al*., 2004; Libault *et al*., 2010; Lohar *et al*., 2006). Additionally, increased phosphorylation of *L. japonicus* MAPK proteins has been observed following inoculation with *M. loti* (Lopez-Gomez *et al*., 2012) and brief induction of ROS has been reported after treating *M. alfafa* with *S. meliloti* (Santos *et al*., 2001), *Phaseolus vulgaris* with Nod factor purified from *Rhizobium etli* (Cárdenas *et al*., 2008) and after treatment of peanut (*Arachis hypogaea*) with *Bradyrhizobium sp* (Munoz *et al*., 2015). The exact mechanism by which rhizobia activate host immunity is currently unknown, but it may not be simply related to activation of MTI by MAMPs; for instance *L. japonicus* has been demonstrated to be sensitive to flg22 but purified flagellin from its symbiont, *M. loti* does not induce immunity (Lopez-Gomez *et al*., 2012). Additionally, treatment with LPS of *S. meliloti* actually reduces ROS production and the expression of genes associated with defence responses in *M. truncatula* callus cell cultures (Tellstrom *et al*., 2007).

The relatively brief nature, typically lasting only for a few hours, of activation of legume immunity by rhizobia is suggestive of active and long-lasting suppression of defence responses. Both host and rhizobial factors have been implicated in moderating defence responses during this symbiosis. EPS of rhizobial cells have been implicated in negatively regulating defence responses and this is a common activity of EPS throughout the bacterial kingdom, either by binding receptors or masking antigens (Fraysse *et al*., 2003). Previous work has demonstrated increased expression of *M. truncatula* defence genes in response to *S. meliloti* mutants deficient in succinoglycan (an EPS) relative to the wild type bacteria (Jones *et al*., 2008). LPS of *Sinorhizobium meliloti* and more specifically, the Lipid A epitope of this is able to suppress ROS production in *M. truncatula*. Curiously, this epitope actually stimulates oxidative burst in the non-legume *N. tabacum* (Scheidle *et al*., 2005; Tellstrom *et al*., 2007). There is evidence to suggest that the Nod factors secreted into the rhizosphere by nodule bacteria can suppress MTI; *G. max* demonstrates reduced ROS production in response to flg22 following pre-treatment with Nod factor from *B. japonicum* (Liang *et al*., 2013). Host factors

are also known to be important in moderating defence responses; *P. sativum* mutants in key symbiosis genes have been demonstrated to produce more robust defence responses after inoculation with *Rhizboium leguminosarum*, suggesting these genes may be important in suppression of immunity and that in their absence, rhizobia are seen as pathogens not symbionts (Ivanova *et al*., 2015). Taken together, there is a strong case that legume and rhizobial effectors have co-evolved such that they are able to selectively minimize the extent of defence responses harmful to symbiosis (Oldroyd, 2013).

***Ralstonia solanacearum* is a broad-spectrum phytopathogen associated with characteristic wilting in a wide variety of plant species:** *Ralstonia solanacearum* is a bacterial pathogen found in soil that is able to infect and cause disease in over 200 different plant species including commercially important crops such as potato and tomato (Salanoubat *et al*., 2002). The combination of its exceptionally large host range, persistence in the environment (*R. solanacearum* can survive in many soil types for months without invading a host plant (van Elsas *et al*., 2000)) and the lethality of the disease it is associated with (bacterial wilt) mean *R. solanacearum* is one of the most intensely studied plant pathogens.

Molecular determinants contributing to *R. solanacearum* pathogenicity: *R. solanacearum* has been found to secrete a large number of proteins through type II and III secretion systems. Many of these proteins are important or essential for virulence, suggesting that the pathogen relies on using effectors to manipulate the host by either down-regulating defence pathways or enhancing access to nutrients (Poueymiro and Genin, 2009).

The pathogenesis cycle of *R. solanacearum* begins when bacteria in the soil recognize the presence of a potential host plant nearby. This is thought to be mediated by PrhA on the bacterial membrane which is directed against some of cell wall pattern and activates expression of virulence genes upon its detection (Aldon *et al*., 2000). Determinants of motility (*fliC* and *fliM*) and chemotaxis (*CheA* and *CheW*) are also important virulence factors for the bacteria, suggesting they have roles in mediating contact between the pathogen and host (Tans-Kersten *et al*., 2001; Yao and Allen, 2006).

The bacteria then gain entry to the plant through naturally occurring gaps in the roots or wounds. The bacteria has been observed to preferentially colonize the elongation zone of the primary root and the primordia or emerging lateral roots where the epidermis is weakest (Vasse *et al*., 1995). The bacteria is also known to possess several hydrolytic enzymes that are capable of degrading components of plant cell walls, which are thus contributors to virulence, although it is not known if these contribute to the capacity to invade the root directly (Huang and Allen, 1997; Schell, 1987; Tans-Kersten *et al*., 1998).

Once a bacterium has entered the root it then begins to divide and colonize the intracellular spaces of the underlying cortex. From here, there is penetration of the vasculature and ultimately invasion of the shoots via the xylem (Vasse *et al*., 1995). This would involve breaching the Casparian strip of the endodermis and the mechanism by which this is achieved is unclear. The bacterium possesses two putative drug efflux pumps, *acrA* and *dinF*, which likely protect the pathogen from host antimicrobials during the colonization process (Brown *et al*., 2007).

The bulk of *R. solanacearum* EPS is comprised of the exopolysaccharide EPS I and this is thought to be only produced when bacterial densities in the xylem are sufficient via a quorum sensing mechanism (Lowe-Power *et al*., 2018). Excessive secretion of EPS by the bacteria is thought to facilitate the blockage of the vasculature and be responsible for the characteristic wilting and death of infected plants (Kao *et al*., 1992).

***R. solanacearum* as a model for legume-pathogen interactions:** Although legumes are not commonly used as a model to study pathogenesis in plants, the interaction with *R. solanacearum* represents the best characterized *M truncatula* pathosystem and best characterized pathosystem involving legumes. Some strains of *R. solanacearum* are able to induce disease symptoms in the susceptible *M. truncatula* A17 wild type, cumulating in characteristic wilting in a matter of days to weeks (Vailleau *et al*., 2007). Root pathogenesis in *M. truncatula* is dependent on the activity of Type III Secretion Systems possessed by the bacteria (Turner *et al*., 2009) and one particular effector, GALA7, is essential for infection (Angot *et al*., 2006). *R.*

solanacearum has also been demonstrated to cause limited pathogenesis in the legume *L. japonicus* although this study used inoculated leaves rather than roots (Nagata *et al*., 2008).

Moreau *et al*. has addressed the transcriptional response of the *M. truncatula* to *R. solanacearum* and *S. meliloti* and identified the TF *EFD* to be strongly up-regulated in response to both. Furthermore, *efd-1* null mutant plants mutants are resistant to infection by *R. solanacearum,* suggesting that this host gene contributes virulence. The positive effect of *EFD* pathogenesis was dependent on cytokinin signalling (Moreau *et al*., 2014). *MtEFD* is also known to be up-regulated during nodulation and promote nodule maturation at the expense of the formation of new nodules, an activity also mediated via moderation of cytokinin signalling (Vernie *et al*., 2008). Although miRNAs involved in legume symbiosis have been researched extensively, there has been very little published research on the role of miRNAs in defence response responses in legumes. Formey *et al*. have identified miRNAs responsive to *R. solanacearum* and predicted targets of these, but no functional analysis has taken place beyond this (Formey *et al*., 2014). QTL analysis has mapped the basis of resistance to *R. solanacearum* infection to the *MtQRRS1* locus which is predicted to comprise 7 R genes (Ben *et al*., 2013), but little else is known about the locus besides this.

In investigating pathogen-symbiosis intersections, a synthetic biology approach using a strain of *R. solanacearum* transformed with the symbiotic plasmid from *Cupriavidus taiwanensis* (the natural symbiont of the legume *Mimosa pudica*) was carried out. Interestingly, this work has shown that a small number of mutations is sufficient to allow a non-rhizobial strain of bacteria to induce some stages of nodulation. It should be noted that Ralstonia is not able to cause disease in *M*. *pudica*, but inactivation of the gene *hrcV* (a structural component of a type III secretion system) resulted in the formation of nodule-like structures and knockout of *hrpG* allowed colonization of nodule cells by the bacteria (Marchetti *et al*., 2010). The nodule-like structures formed by the bacteria resembled nodules in appearance and were invaded by infection threads. Subsequent experiments using the same system led to reduced immunogenicity of the bacteria (Marchetti *et al*., 2017; Marchetti *et al*., 2014). Although this cannot be considered a mutualism because the

resulting nodules were unable to fix nitrogen, this demonstrates a pathway that a pathogen could transition into a commensal or at least non-pathogenic invader, in line with theory that nodulation evolved from parasitic interactions.

Aims: immunity is important during symbiotic interactions between legumes and rhizobia with moderation of host immunity usually required for successful nodulation. The current hypothesis is that during the initial interaction between the bacteria and plant, rhizobial MAMPs initially activate MTI but this is quickly supressed and the symbiosis is able to proceed (Cao *et al*., 2017). The molecular determinants of this transition, especially from the host plant perspective, are poorly determined. Therefore, the response of the model legume *M. truncatula* to its symbiont *Sinorhizobium medicae* and the pathogen *Ralstonia solanacearum* was contrasted over time to identify genes that could be responsible for the transition from a state of defence to symbiosis. This was carried out by first using RT-qPCR to assay timings of response using genes that serve as markers for defence and symbiosis. Subsequently, RNAseq was used to identify novel transcripts involved in these processes.

Results:

To maximize physiological relevance plants used in these experiments (Figure 4.1) was grown in pots containing perlite and vermiculite (Figure 4.1A), rather than on agar plates. Perlite and vermiculite mimic the physical properties of soil but can be autoclaved which permitted these experiments to be conducted in pseudosterile conditions. The system also allowed seedlings sufficient space for root growth over an extended period of time, enabling mature roots to be studied and sampled.

Plants were inoculated with either a symbiont or pathogen. *Sinorhizobium medicae* WSM419 (‘rhizobia’, Figure 4.1D) was used to induce nodulation in plants instead of the model *Sinorhizobium meliloti* 1021 system as the former is a naturally occurring symbiont of *M. truncatula* and the resulting symbiosis is more efficient in terms or nitrogen fixation (Terpolilli *et al*., 2008). *R. solanacearum* (‘ralstonia’, Figure 4.1E) was used as a pathogen (see below for rationale). Plants infected with ralstonia were grown with replete

nitrogen (15 mM NH_4NO_3) to ensure they were sufficiently resourced to form defence responses whilst plants inoculated with rhizobia were grown in the complete absence of nitrogen to avoid inhibition of nodulation. Therefore, control plants were also grown in the presence (Figure 4.1C) or absence (Figure 4.1B) of nitrogen to serve as baselines for these experiments. At intervals (Figure 4.1F) chosen to encompass early interactions with the treatments up to late stage ralstonia pathogenesis and nodules being visible on rhizobia-treated roots (Figure 4.1G), the whole root system, including nodules were applicable, was harvested (Figure 4.1H-K, see methods), RNA extracted and gene expression analysis carried out using RT-qPCR and RNAseq.

***R. solanacearum* is able to infect and induce wilting in *M. truncatula*:** To compare defence and symbiosis responses in plants, a biotic treatment capable of inducing the former was required. For this purpose, the broad-spectrum root pathogen *R. solanacearum* was tested. To clarify that ralstonia is able to infect and cause disease in *M. truncatula* within the conditions of these experiments, 3 day old seedlings were treated with ralstonia in the same perlite and vermiculite pot system used in subsequent experiments (see methods) and observed for symptoms of infection. Initially, infected plants do not show any signs of disease but then plants begin to wilt dramatically around 2 weeks after infection with apparent death of the host plant occurring 3-4 weeks after the initial inoculation (Figure 4.2). The characteristic wilting symptoms and timescale of this event is broadly consistent with ralstonia infection in other susceptible species (Mukhtar *et al*., 2008; Tans-Kersten *et al*., 2001).

To clarify the efficacy of *S. medicae* strain 419 in the perlite-vermiculite pot system, plants were also grown for 4 weeks following *S. medicae* inoculation. Plants inoculated with rhizobia (Figure 4.3C) had observably more above ground biomass than mock treated plants (Figure 4.3A) although by this stage, the total amount of nitrogen fixed was not sufficient to catch up with the size of plants grown with replete nitrogen (Figure 4.3B).

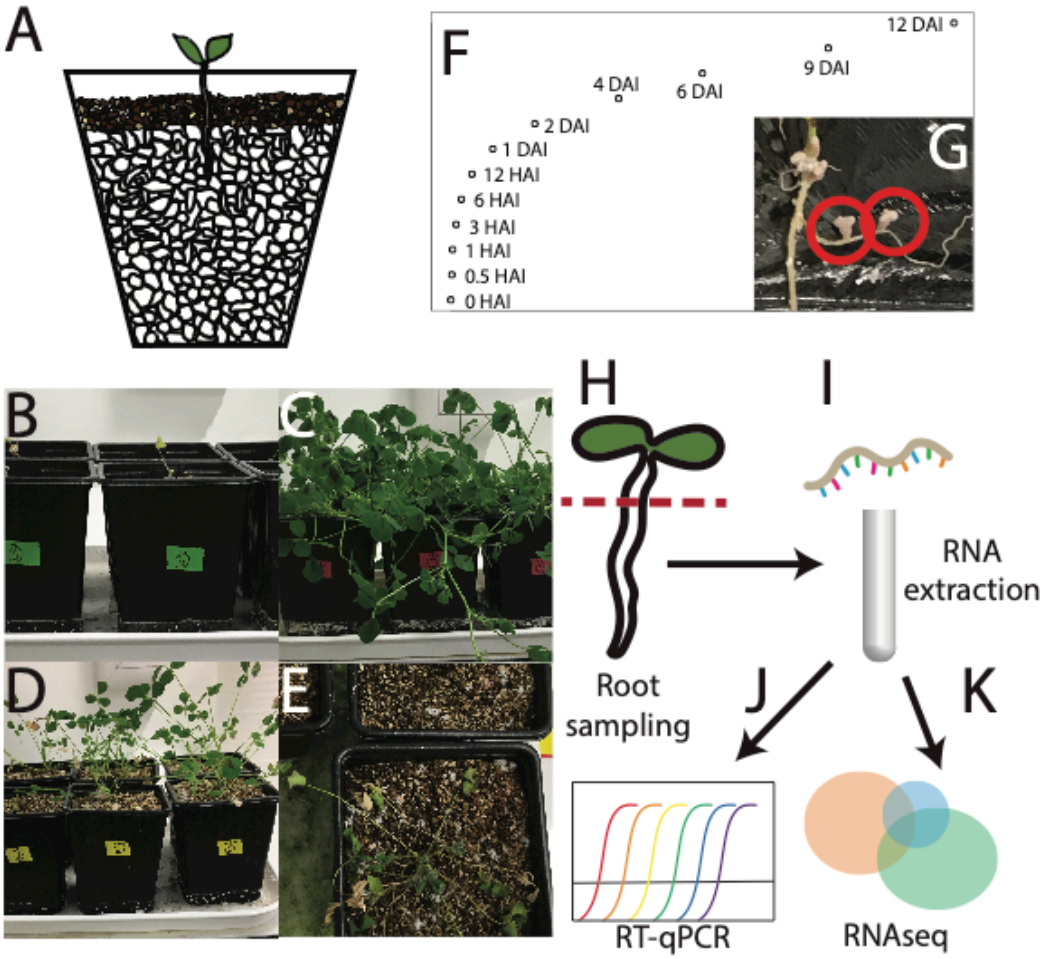

Figure 4.1: Overview of experimental design. 2 day old *M. truncatula* seedlings were transplanted into pots (A) filled with perlite and vermiculite and watered from below with a nutrient solution containing 0 (B) or 15 mM ammonium nitrate (C). Within each nitrogen level, plants were either mock-inoculated or inoculated with *S. medicae* (D) or *R. solanacearum* (E) respectively. At a range of timepoints (F) from immediately before to 12 days after inoculation, by which point nodules were visible on rhizobia treated roots (G), the whole root system of seedlings were harvested (H) and subject to RNA extraction (I). Samples were than assayed (J) for the expression of putative marker genes for symbiosis or defence by RT-qPCR whilst select samples were analysed using RNAseq, based on RT-qPCR results (K).

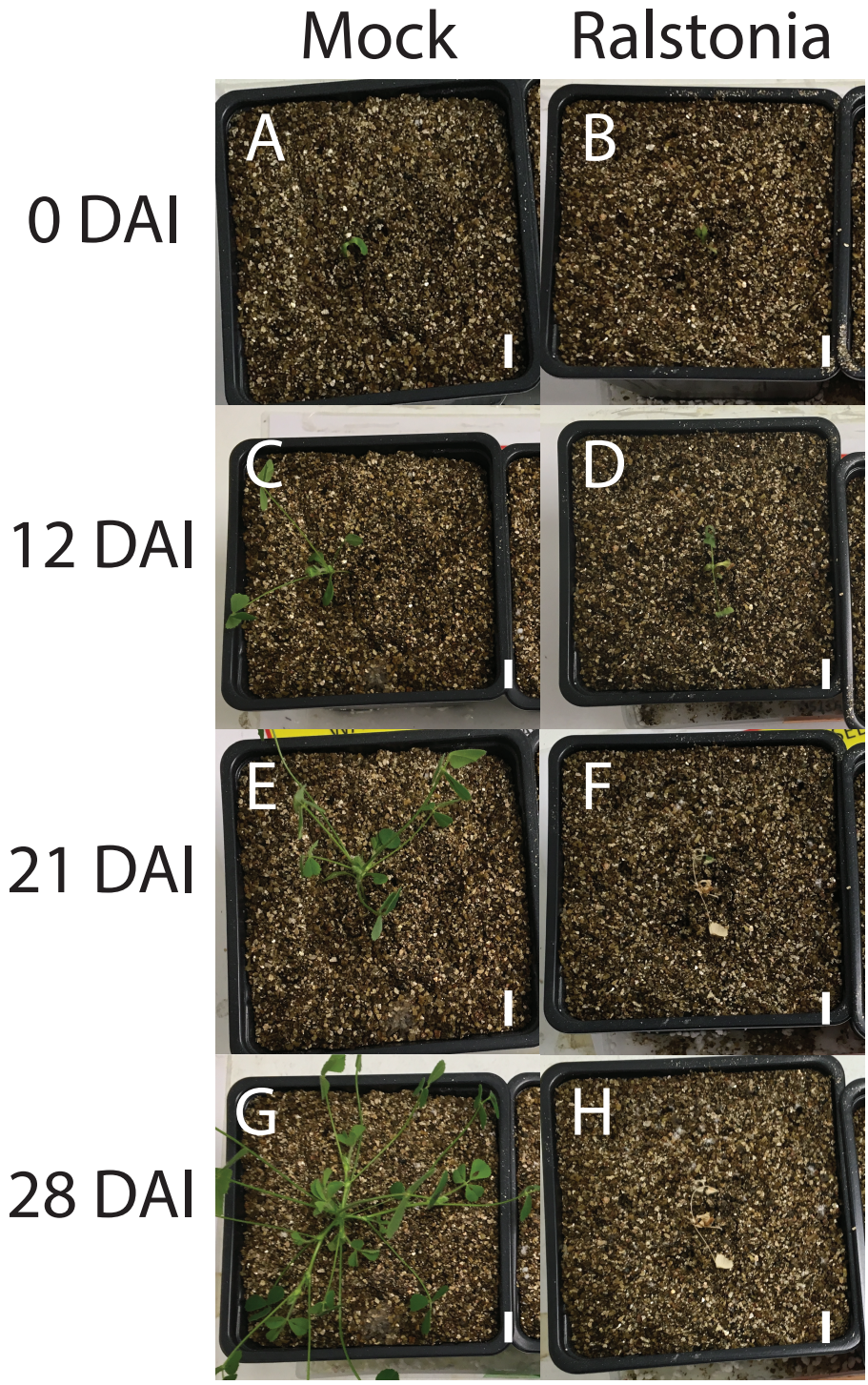

Figure 4.2: Phenotype of *M truncatula* plants infected with *R. solanacearum*. Representative picture of plants either mock-inoculated (left) or inoculated with *R. solanacearum* (ralstonia, right) (B) photographed at immediately before (A/B) 12 (C/D), 21 (E/F) and 28 (G/H) days after inoculation (DAI). Scale bars represent 1 cm. n=6.

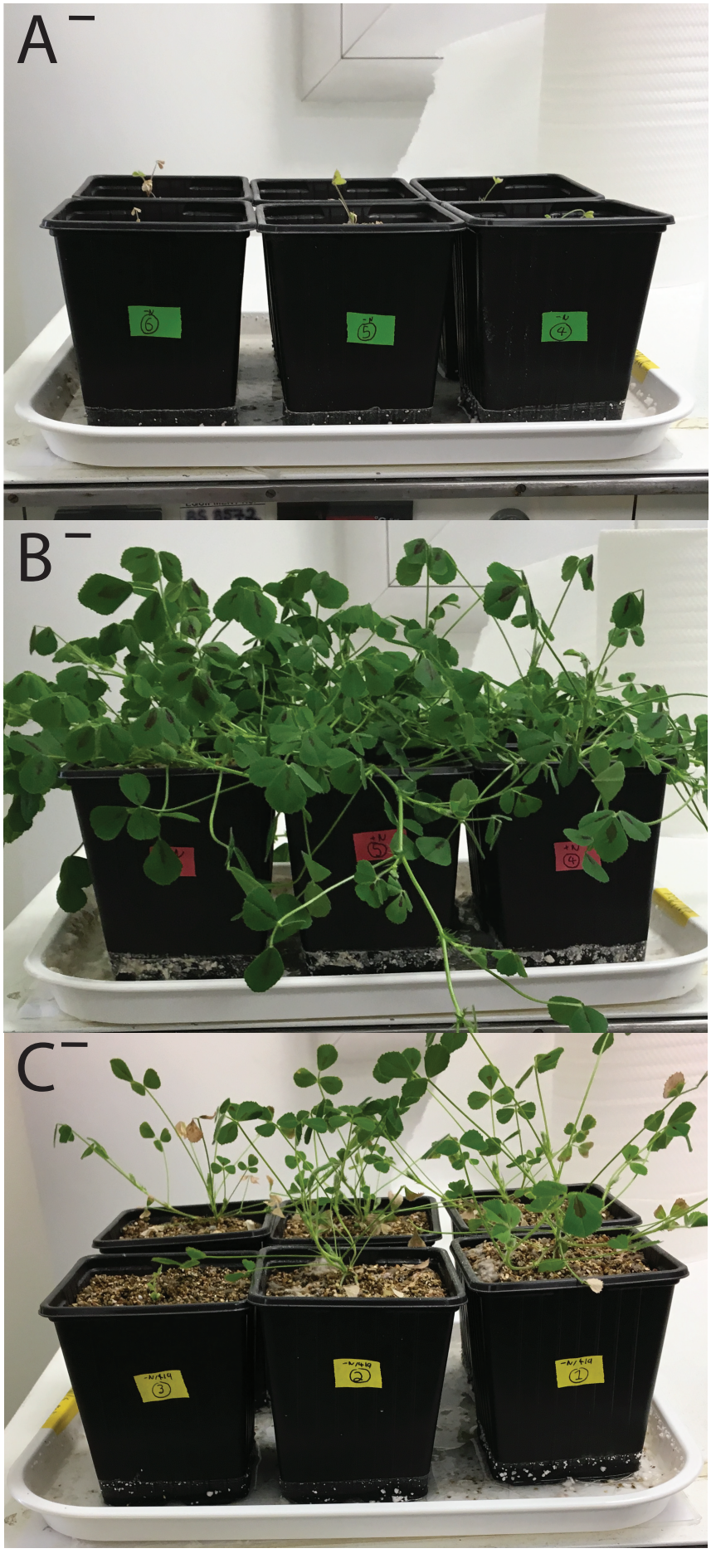
A⁻
B⁻
C⁻

Figure 4.3: Phenotypic effects of nitrogen treatment and *S. medicae* inoculation on growth in *M. truncatula*. Representative plants photographed after 4 weeks growth with no nitrogen (A), 15 mM ammonium nitrate (B) or no nitrogen and inoculated with *S. medicae* 419. Scale bars (to right of panel labels) represent 1 cm. n=6.

Identifying marker genes for symbiosis and defence:

To identify at which timepoints the response to rhizobia and ralstonia occurred and the timescale of the response, RT-qPCR was employed to assay the expression of marker genes that are associated with nodulation and defence responses in *M. truncatula*. Therefore, approaches were taken to identify candidate genes that may function as markers for these processes.

ENOD11 (Early Nodulin 11, Medtr3g415670), is commonly used as a marker for symbiosis in *M. truncatula* as is up-regulated during interactions with both rhizobia and arbuscular mycorrhizae (Journet *et al*., 2001). *NIN* (Medtr5g099060) is responsible for regulating many events during nodulation and is also known to be strongly induced by rhizobia treatment (Vernie *et al*., 2015). These two genes were chosen as symbiosis markers.

Although marker genes for responses to rhizobial symbiosis in *M. truncatula* are well characterized, the pathways of defence and are not so well established in this model. Therefore, one approach to identify marker genes for these responses is to identify homologs from the corresponding and much better studied pathway in *A. thaliana*. To identify homologs of *A. thaliana* genes in *M. truncatula*, the protein homologs tool on Phytozome v12.1 (Goodstein *et al*., 2012) was used to query each gene of interest using the sequence from the *A. thaliana* TAIR10 annotation version (Berardini *et al*., 2015; Huala *et al*., 2001) against sequences from the *M. truncatula* Mt4.0 annotation version (Tang *et al*., 2014), using protein similarity as a metric. Candidate homologs were determined on the basis of the extent of similarity between the protein sequences. In instances where multiple genes very closely matched the candidate (to be expected due to the genome duplication of *M. truncatula* (Young *et al*., 2011)) all hits were considered as potential homologs and were considered for further study.

Candidate marker genes for defence responses were selected from genes known to be involved in the *A. thaliana* innate immune pathway. *FLS2* (At5g46330) is a receptor-like kinase that is sensitive to the MAMP flagellin commonly found on many bacteria and to flg22, a 22 amino acid epitope within this (Chinchilla *et al*., 2006; Gomez-Gomez and Boller, 2000). *CYP82C2* (Cytochrome P450, family 82, subfamily C, polypetide 2, At4g31970), *PHI1* (Phosphate-induced 1, At1g35140) and *FOX* (FAD-linked oxidoreductase, At1g26380) are all strongly up-regulated by flg22 (Boudsocq *et al*., 2010). The enzyme-encoding gene *ICS1* (Isochorismate synthase 1, At1g74710) is responsible for the synthesis of the defence hormone SA during plant defence responses and is also used as a defence marker (Wildermuth *et al*., 2001).

In addition, a literature review of the candidate homologs was used to provide further evidence in support of their use as markers for nodulation or defence. The top hit for *FLS2* (Medtr4g094610) was found to have been studied previously in *M*. *truncatula* A17 and had been demonstrated to be strongly up-regulated in response to *Pseudomonas syringae* DC3000 (Chen *et al*., 2017). This supported the prediction that this gene has the same function as its *A*. *thaliana* counterpart and may therefore function as good a marker for defence responses triggered by flg22 or *R. solanacearum* inoculation. Researching one putative homolog of a gene ultimately not used in this study showed it had been used as a defence marker in *M. truncatula* in response to the pathogen *Phytophthora palmivora* (Rey *et al*., 2017). Another gene used as a defence marker in this same study, Medtr1g075340, was included.

Once candidate genes were identified, transcript sequences were recovered from the *M*. *truncatula* genome annotation 4.0v2. Primers were designed against these sequences using Primer3 (Koressaar and Remm, 2007; Untergasser *et al*., 2012). Once primers were designed they were aligned against the *M*. *truncatula* genome using BLAST (Altschul *et al*., 1990) to ensure they only bound within the sequence of the gene they were directed against. For full list of primers used, see Table 4.1. Primers were used for the genes *Mt*Ubiquitin, *Mt*UBC9 and *Mt*β-tubulin (Kakar *et al*., 2008) to determine housekeeping gene expression and the average of the expression of the three genes was used to normalize marker gene expression.

Target gene	Identifier	Forward primer	Reverse primer	AL
MtFLS2	Medtr4g094610	TGGAACGGAACAGCTTGTCA	TCAGGATCAGAGAGGGTGCA	114
MtNIN	Medtr5g099060	TTCCAAGAAGCGTGCAGTGA	GTTGGCATGGTGAGCAAGTG	141
MtENOD11	Medtr3g415670	TGGTAACCAGCCTCCACCTA	CCACATGCAAAGATGGGACG	210
MtFOX	Medtr2g031590	GGGAGCAACTGACACTAGCA	CAGAAAGAGGCGATGCTGGA	212
MtCYP82C	Medtr6g008530	ACATGGTCGTTTCGTCTCGT	ACAATCTTGCGAAGTTGGCG	115
MtICS1	Medtr4g076520	TGTTGGTTCGGCTGTTTCCT	ACCTCCATATGCGCGAATCA	112
MtPHI1	Medtr5g464340	TCGCGTACATTTGGGTTGGA	CCATTCCGTCAAGACCCACA	135
Unknown	Medtr1g075340	AGTTGGTTGGAGTGGAAGGAA	TCCAAACCCAAACCCTGTCA	112
MtUbiquitin	Medtr3g091400	GCAGATAGACACGCTGGGA	AACTCTTGGGCAGGCAATAA	100
MtUBC9	Medtr7g116940	GGTTGATTGCTCTTCTCTCCCC	AAGTGATTGCTCGTCCAACCC	100
Mtβ-tubulin	Medtr7g089120	TTTGCTCCTCTTACATCCCGTG	GCAGCACACATCATGTTTTTGG	100

Table 4.1: Primer sequences used for RT-qPCR analysis. Primers were diluted to a working concentration of 10 µM and stored at -20 C° prior to usage. AL = Amplicon length.

To validate the success of the rhizobial inoculation and deduce which samples show the strongest response to rhizobia, indicative of nodulation establishment, the expression of *ENOD11* (Figure 4.4A) and *NIN* (Figure 4.4B) in roots infected with rhizobia was tested by RT-qPCR. Both genes behave similarly with very low basal expression in uninfected roots. *ENOD11* first has significant up-regulation at 4 DAI with rhizobia and is maintained at greatly elevated expression until the final 12 DAI timepoint. *NIN* expression is not significantly higher upon rhizobial inoculation until 2 DAI and then this level of expression is maintained until the end of the experiment. The timing and amplitude of these expression changes are broadly consistent with the behaviour of these genes upon inoculation with other strains of rhizobia (Journet *et al*., 2001; Vernie *et al*., 2015).

Next, to test if ralstonia or rhizobia inoculation induced transcription of putative *FLS2* homolog Medtr4g094610, expression of this gene was tested by RT-qPCR (Figure 4.5). In rhizobia-inoculated plants, *MtFLS2* shows significantly increased expression at 2 DAI (Figure 4.5A) but not at other timepoints. *MtFLS2* does not show significant DE at any point relative to mock inoculated plants after ralstonia inoculation (Figure 4.5B).

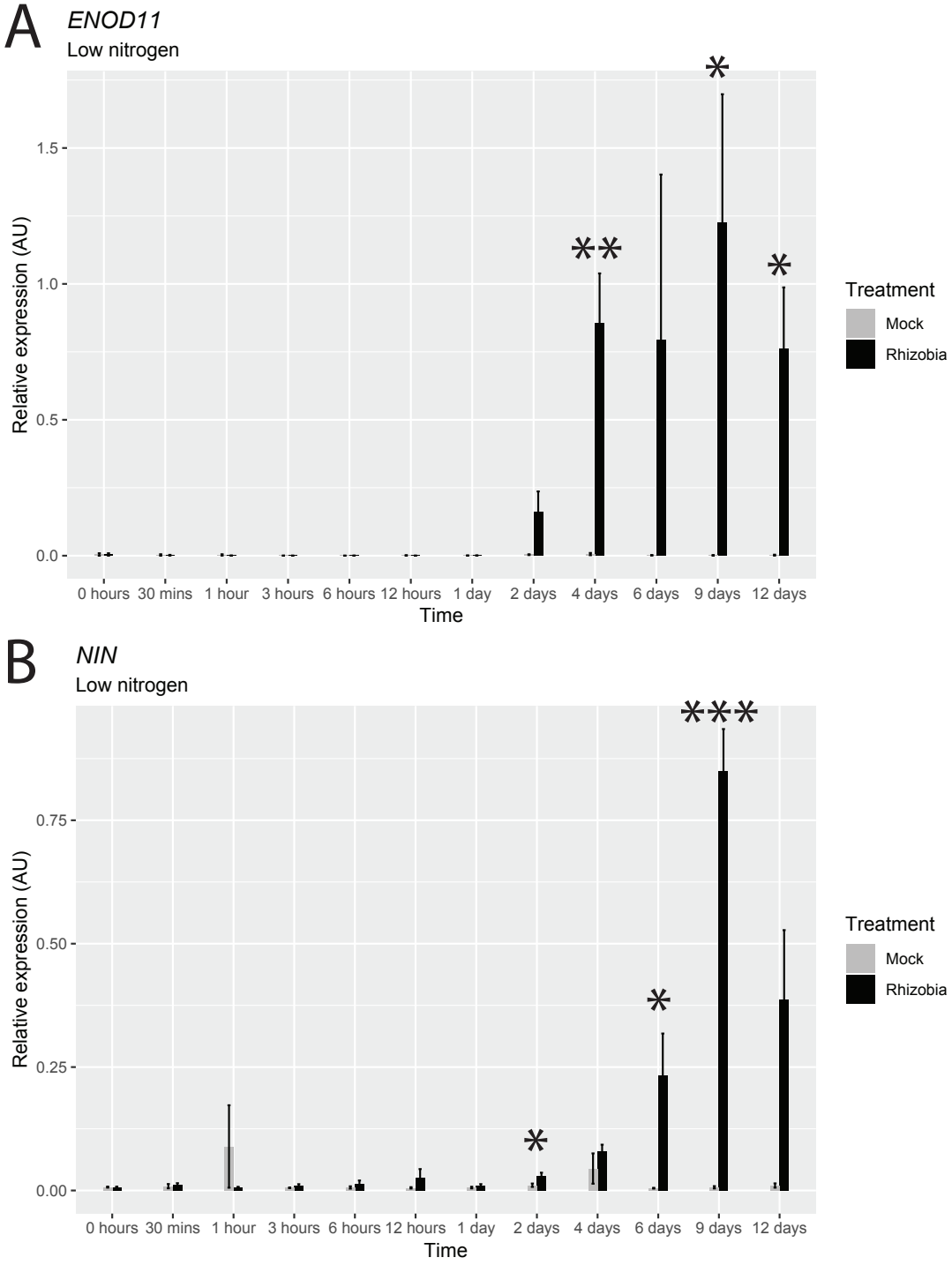

Figure 4.4: Expression of symbiosis marker genes *NIN* and *ENOD11* in *M. truncatula* roots. Expression of *ENOD11* (A) and *NIN* (B) relative to average expression of housekeeping genes Medtr3g091400, Medtr7g116940 and Medtr7g089120 in mock- and rhizobia-treated roots. Error bars represent standard error (n=3+ biological replicates each consisting of 5 seedlings) and asterisks denote significant differences compared to mock-inoculated plants as determined by the Student's T-test at $P < 0.05$ (*), 0.01 (**) or 0.001 (***).

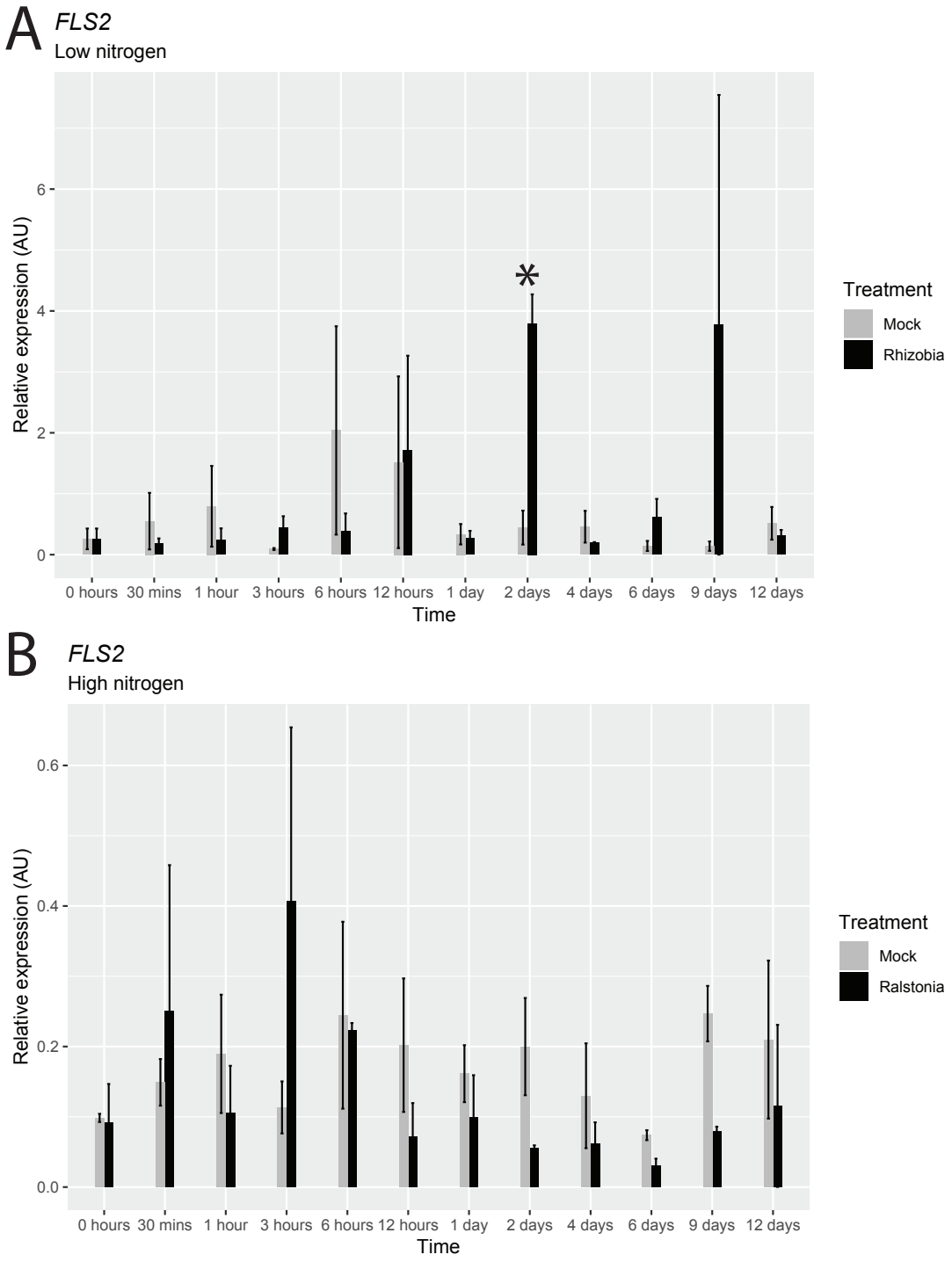

Figure 4.5: Expression of putative *FLS2* homolog Medtr4g094610 in *M. truncatula* roots. Expression of *MtFLS2* in response to the pathogen *R. solanacearum* (A) and the symbiont *S. medicae* (B) relative to average expression of housekeeping genes Medtr3g091400, Medtr7g116940 and Medtr7g089120. Error bars represent standard error (n=2 biological replicates each consisting of 5 seedlings) and asterisks denote significant differences

compared to mock-inoculated plants as determined by the Student's T-test at $P < 0.05$ (*).

As shown above, *FLS2* expression was not induced by ralstonia as would have been expected. To attempt to identify alternative defence marker genes but also to explore the response of *Mt*FLS2, RT-qPCR was used to assay putative defence marker gene expression in plants treated with 100 nM of the immunogenic peptide flg22 (Figure 4.6). This concentration was chosen because it has been shown to elicit defence responses in the model plant *A. thaliana* (Boudsocq *et al*., 2010). Because transcriptional responses to flg22 in *A. thaliana* are short-term, this assay was used only for samples from plants at 30 mins, 1 hour and 3 hours after flg22 treatment. The putative *M. truncatula* homologs of 6 genes associated with defence in *A. thaliana*; Medtr1g075340, *FOX*, *CYP82C*, *ICS1*, *PHI-1* and *FLS2* were tested.

None of the genes tested were significantly induced by flg22 relative to the 0H (untreated) timepoint, including *MtFLS2* (Figure 4.6F). Interestingly, the basal expression of some of these genes, as determined at this first timepoint. Defence markers would be expected to exhibit very low basal expression. *MtFOX* (Figure 4.6B), *MtCYP82C* (Figure 4.6C), *MtICS1* (Figure 4.6D) is over tenfold higher than the average expression of the housekeeping genes used to calculate their relative expression.

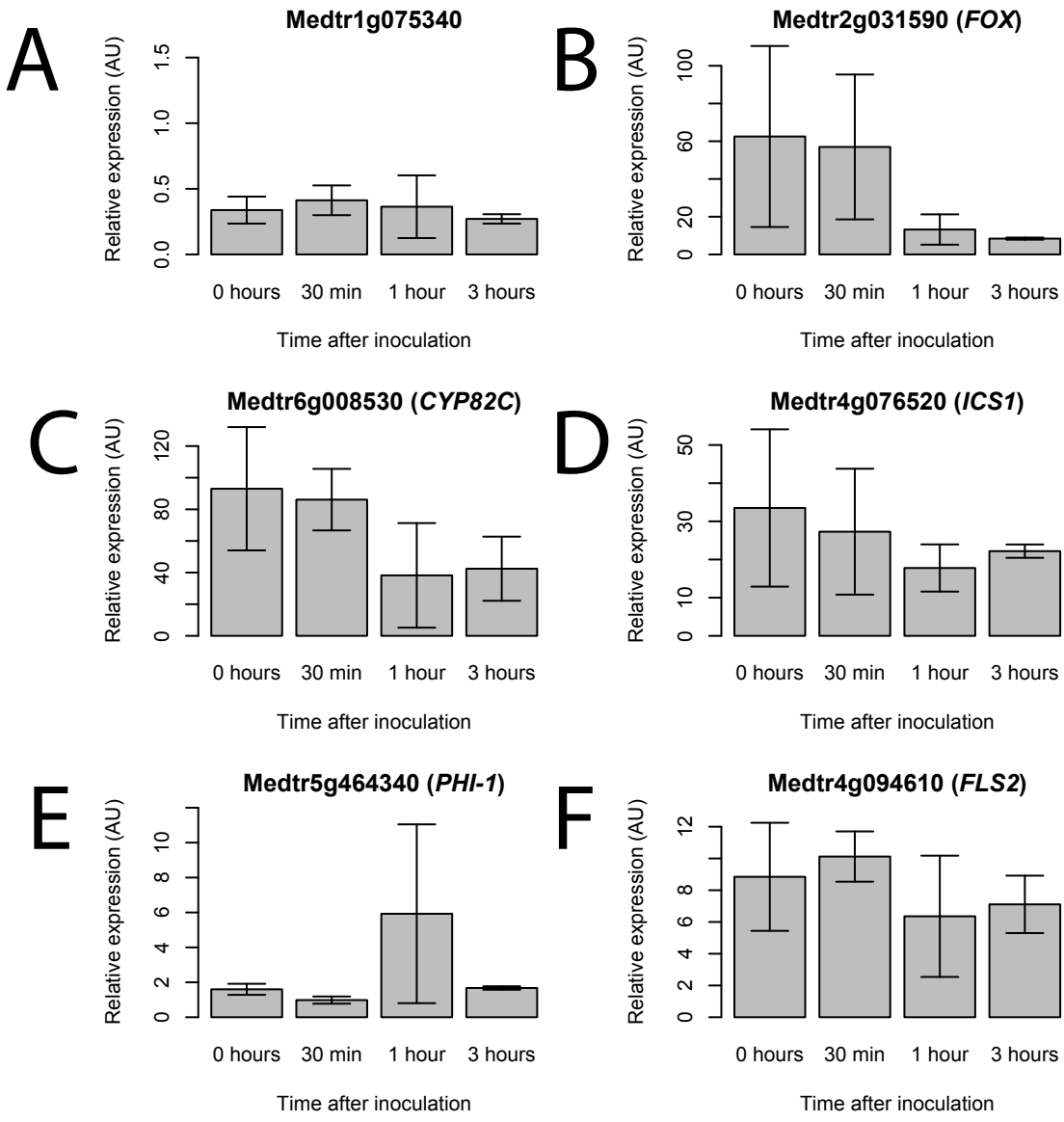

Figure 4.6: Expression of putative defence marker genes in response to flg22. Expression of Medtr1g075340 (A), *MtFOX* (B), *MtCYP82C* (C), *MtICS1* (D), *MtPHI-1* (E) and *MtFLS2* in response to the treatment with 100 nM flg22 relative to average expression of housekeeping genes Medtr3g091400, Medtr7g116940 and Medtr7g089120. Error bars represent standard error (n=2 biological replicates each consisting of 5 seedlings). No significant differences were found compared to plants 0 HAI as determined by the Student's T-test.

Using RNAseq to identify rhizobia- and ralstonia-responsive genes: Overall, the results from RT-qPCR were inconclusive, but longer-term effects of bacterial inoculation suggested that the RNA samples could be informative for further study. To identify genes responding to ralstonia, to be used as

possible markers of *M. truncatula* defence responses in a comparison of, the regulation of defence and symbiotic responses in this model, select samples were used for whole transcriptome profiling using RNAseq. The samples chosen were rhizobia and ralstonia 3 and 12 hours after inoculation, in addition to samples taken immediately before inoculation (0H) in both replete and deplete nitrogen conditions to serve as a reference for the treated samples. The 3 and 12 hour timepoints were chosen to cover an early stage of the interaction between ralstonia and the plant, based on the hypothesis that these were more likely to at the point of MTI in the former case, and at a stage which ralstonia may be using effectors to reprogram host cells and likely induce ETI in the latter case.

Libraries were prepared and sequenced with Illumina sequencing (see methods). Reads were trimmed and mapped to the *M. truncatula* genome annotation 4.0v2. Across all samples 84-88% of raw reads were uniquely mapped following processing representing between 16.5 and 30 million uniquely mapped reads. Principal component analysis was first used to ask if there were differences between treatments (Figure 4.7). The majority of treatments clustered tightly together although the first two principal components displayed on the plot only account for just over half of the total variance in the dataset. Both rhizobia and ralstonia 12 HAI samples are noticeably separated from the rest of the samples on the basis of the first two principal components.

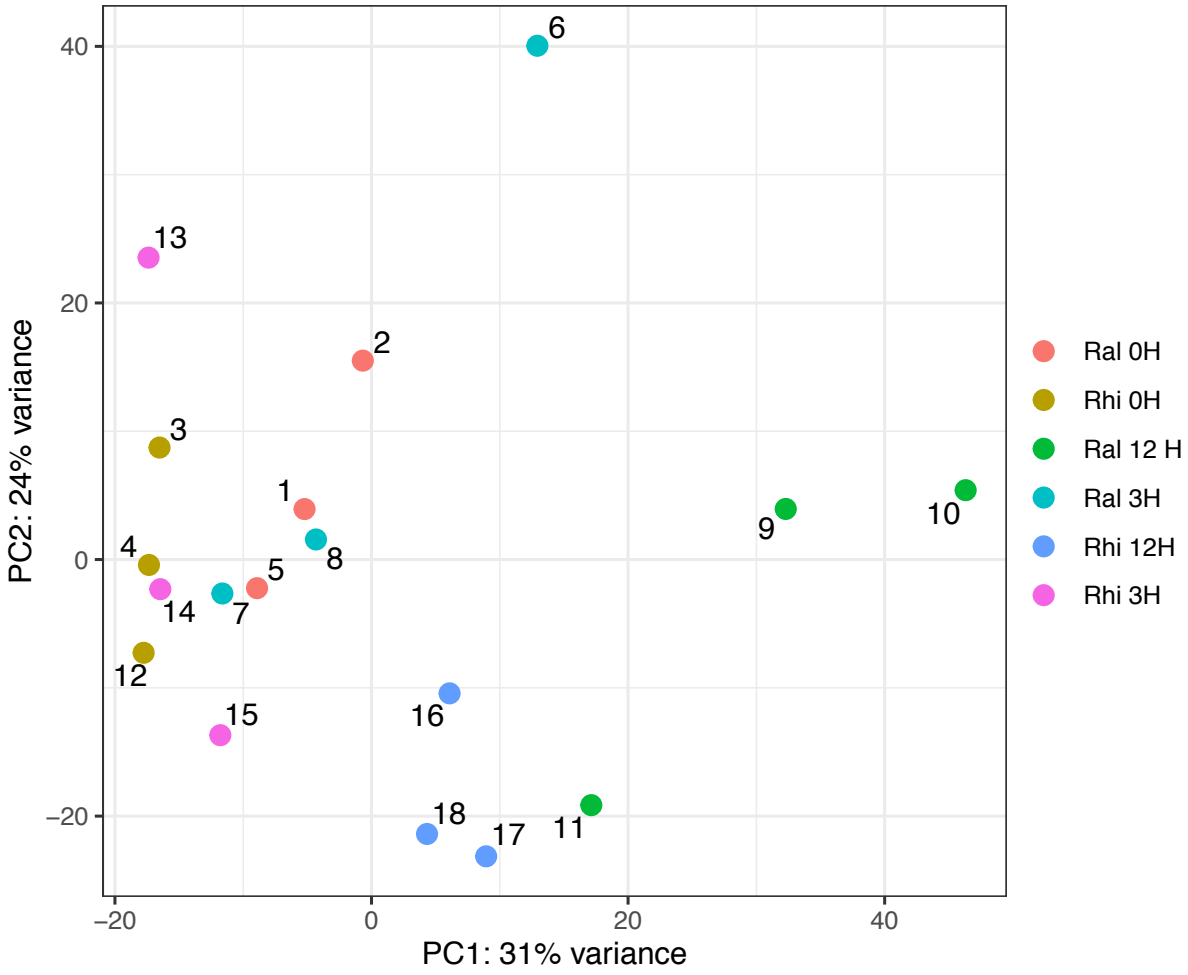

Figure 4.7: Principal component analysis of samples used in RNAseq experiments. PCA plot of normalized RNAseq aligned counts data from samples treated with *R. solanacearum* immediately before (red), 3 hours (cyan) and 12 hours (green) after inoculation or *S. medicae* immediately before (gold), 3 hours (magenta) or 12 hours (blue) after inoculation.

Next, DE genes were calculated using DEseq2 (Table 4.2) in the 3 HAI and 12 HAI rhizobia and ralstonia samples relative to the 0 HAI (immediately before inoculation) samples (Figure 4.8A). The number of genes showing significant changes in regulation in the 3 HAI samples is small with 52 (44 up-regulated, 8 down-regulated) and 42 (22 up-regulated, 20 down-regulated)

genes DE in the rhizobia and ralstonia samples respectively and increases to 222 (158 up-regulated, 64 down-regulated) and 622 (413 up-regulated, 209 down-regulated) in the corresponding 12 HAI samples.

Experiment	Rhizobia 3 HAI	Rhizobia 12 HAI	Ralstonia 3 HAI	Ralstonia 12 HAI
Total DEGs	52	222	42	622
Induced	44	158	22	313
Repressed	8	64	20	209

Table 4.2: Differential gene expression induced by ralstonia or rhizobia inoculation.

Although a proportion of DE genes are unique to a single experiment, over half of the genes DE in every treatment except ralstonia 12 HAI are also DE in other experiments (Figure 4.8B). This is most likely due to conservation of gene expression responses between timepoints of the same inoculant (i.e. rhizobia 3HAI and 12 HAI) and as a consequence of circadian effects (see discussion).

There are three genes which are differentially expressed in every single experiment. Two of these genes, annotated to be a cation transport ATPase (Medtr8g013810) and a spermidine hydroxycinnamoyl transferase (Medtr3g049150) are amongst the most strongly up-regulated in response to both ralstonia and rhizobia, at both timepoints, indicating either strong early increase related to circadian regulation or an important role in both ralstonia and rhizobia responses. The third, Medtr2g024300 is very strongly down-regulated in both rhizobia and ralstonia 3 HAI but up-regulated in response to both treatments 12 HAI. This pattern is more consistent with regulation according to a circadian pattern and thus it might not be directly related to ralstonia or rhizobia responses. Another likely indicator of circadian effects in the dataset is that all 5 genes that are DE in both treatments at 3 HAI and all but 2 of the 116 DE genes at 12 HAI show the same direction of change (induction or repression) in both treatments with similar fold changes. Although not present in the WL3 dataset, Medtr1g098160 is upregulated in the other 3

datasets. This protein has high homology to the *A. thaliana* protein *GIGANTEA* which is a circadian regulated gene that peaks in expression around 10 hours after dawn (Fowler *et al*., 1999) suggesting that regulation of this protein is conserved in *M. truncatula*.

Direct comparisons between ralstonia and rhizobia inoculated plants are difficult because ralstonia inoculation was carried out on plants grown in replete N whilst plants that were inoculated with rhizobia were grown in the absence of N. This issue can be partially addressed by identifying DE genes between the two sets of T=0 samples, which allows for some genes differentially expressed in response to N to be identified. This leads to the identification of 90 genes differentially expressed (70 up-regulated and 20 down-regulated) as a result of the differing N levels. These include many genes with annotations suggestive of a role in nitrogen responses (see discussion).

To explore the processes that change upon rhizobia or ralstonia inoculation, a GO term overrepresentation analysis was carried out. In both the 3 HAI samples for rhizobia and ralstonia, the number of DE genes is small and consequently, overrepresented GO terms are generic, although this is also a consequence of many *M. truncatula* genes being poorly functionally characterized. Despite this limitation there were some interesting results. The GO terms ‘developmental process’ (P=1.8e-03) and ‘phosphatase activity’ (P=3.5e-02) both appear in the rhizobia 3 HAI sample whilst genes from the ralstonia 3 HAI sample are enriched for ‘regulation of cellular biosynthetic processes’ (P=1.20e-02).

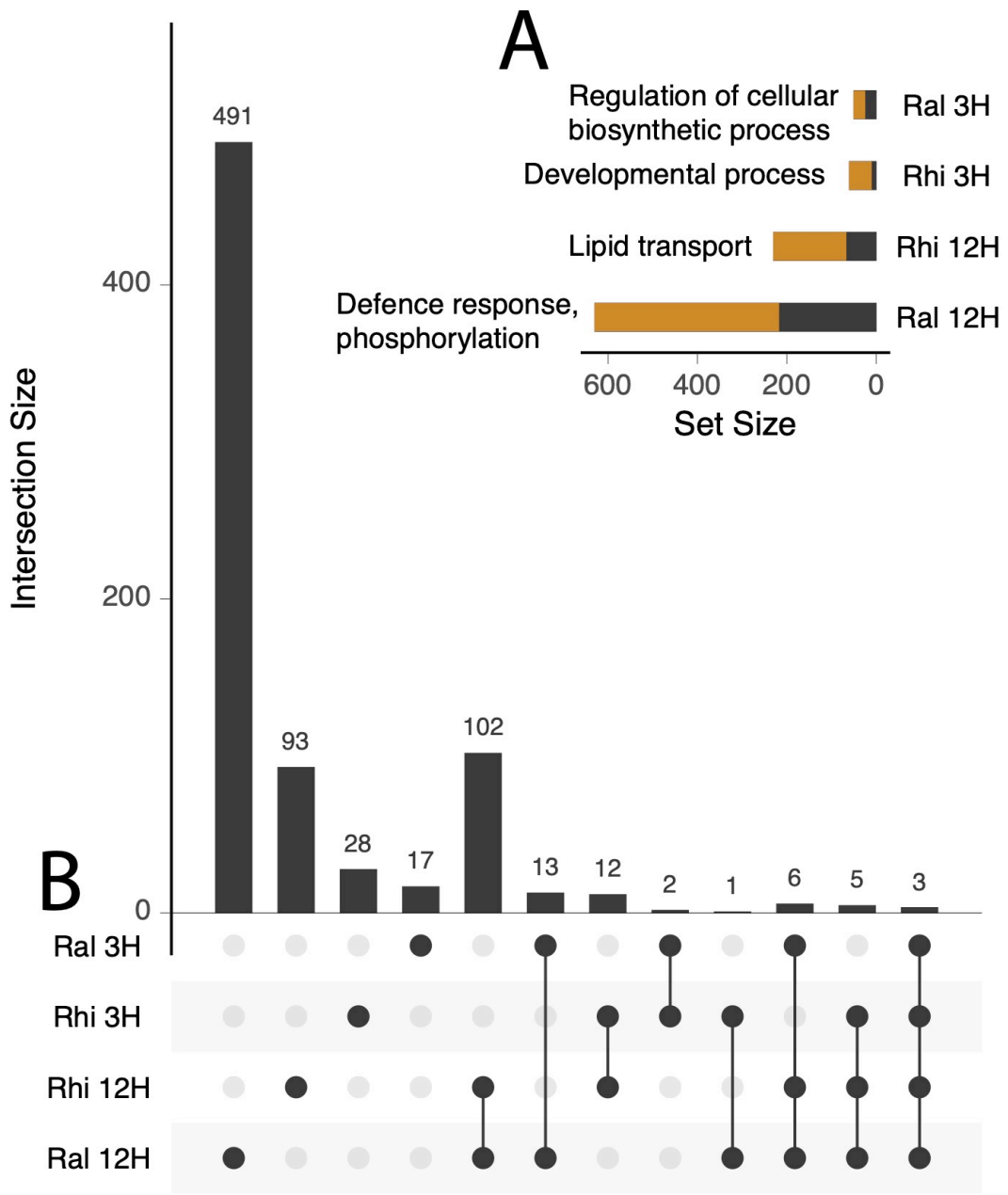

Figure 4.8: Differentially expressed genes in response to *R. solanacearum* and *S. medicae* inoculation 3 and 12 hours after inoculation. Total number of DE genes (A) during each experiment (n=3) with the proportion of these showing up-regulation shaded in orange and those down-regulated in grey. Select GO terms overrepresented amongst experiments (sets) are included. UpSet plot (B) showing overlap between different sets. Where points on the X-axis are connected with lines denotes conservation of DE genes between sets. Comparisons between sets for which there is not any intersection are omitted.

Amongst the 413 genes up-regulated following 12 HAI with ralstonia are 31 genes annotated with the GO term 'defence responses' (P=7.2e-05). 30 of these are predicted to be LRR proteins. LRR proteins typically function as R proteins and therefore require additional proteins to signal through (McHale *et al*., 2006). This would suggest that there are likely additional genes involved in defence pathways but not annotated as such, including targets downstream of these receptors, present in the data which are also involved in defence responses but not annotated as such. The GO term 'phosphorylation' was also overrepresented amongst these 413 genes (P=5.5e-05). Whilst this will likely encompass many generic cellular processes, proteins responsible for signal transduction downstream of LRRs are likely to have kinase activity (McHale *et al*., 2006) and therefore might also be involved in defence responses.

The 158 12 HAI rhizobia induced genes are overrepresented for the GO term 'lipid transport' (P=7.2e-04). Of the 6 genes with this annotation, 4 are repressed and 2 are induced, including the third top hit from the dataset. 8 genes in this datas*et al*so have the annotation 'defence responses' although this is not statistically significant across the set as a whole. 4 of these are shared with the ralstonia 12 HAI set with an additional 2 each annotated as LRRs and defensins.

In conclusion, both rhizobia and ralstonia do induce small numbers of transcriptional changes even as early as 3 HAI and by 12 HAI hundreds of genes are induced or repressed. Many of these genes do not have functional annotations but it is clear from this which are annotated that there is specific responses associated with either bacteria occurring.

Discussion:

***R. solanacearum* is able to induce bacterial wilt in *M. truncatula*:** Previous study by Vailleau *et al*. has demonstrated the susceptibility of *M. truncatula* A17 to *R. solanacearum* (Vailleau *et al*., 2007). Due to differences in experimental design employed in the work presented here, it was first necessary to observe the phenotypic effects of *R. solanacearum* inoculation on plants grown in these conditions. In particular, the age of the seedlings at the point of inoculation were different; seedlings were 3 days old in these

experiments and 13 days old in the Vailleau *et al.* study. Seedlings were inoculated at younger age in the experiments in this chapter as germinating seedlings in the rhizosphere are likely to encounter pathogens and symbionts immediately. Therefore, it was necessary to ensure that not only did *R. solanacearum* still induce disease from this early stage but also that infected plants survived long enough to obtain sufficient root material for molecular analysis.

The symptoms of pathogenesis (Figure 4.2) observed here were identical to those observed in the Vailleau *et al.* study. However, the progression of pathogenesis was slower in these experiments with symptoms of wilt not apparent until 2 weeks after inoculation and wilting occurring 3-4 weeks after inoculation. In the Vailleau *et al.* study, the onset of wilting was observed within a week and the majority of plants were severely wilted or dead within 2 weeks (Vailleau *et al*., 2007). This is likely explained by nutritional differences between the two experiments; plants inoculated with ralstonia in these experiments were watered with a replete nutrient solution and therefore were likely better able to resist the infection.

This clarifies that *R. solanacearum* is a suitable model for studying pathogenesis in *M. truncatula* ecotype A17. The use of ralstonia as a pathosystem allows for defence responses in *M. truncatula*, currently poorly understood, to be studied in greater detail. Understanding how *M. truncatula* responds to pathogens is important because this allows greater insight into the interplay between symbiotic interactions (i.e. nodulation) and pathogenic interactions in the rhizosphere.

***S. medicae* strain WSM419 induces expression of *ENOD11* and *NIN* in *M. truncatula* roots:** *S. meliloti* 1021 is the most commonly used model to study nodulation in *M. truncatula*. However, whilst *M. alfalfa* is a natural host, symbiosis with *M. truncatula* is not a pairing that occurs in nature and its perpetuation as a model in this context is a legacy of it having its genome sequenced prior to other, more compatible symbionts. *S. meliloti*-*M. truncatula* symbiosis does result in nodule formation and some degree of nitrogen fixation but the symbiosis is not optimal (Terpolilli *et al*., 2008). Therefore, the experiments in this chapter were carried out using *S. medicae* 419 which is

able to fix nitrogen with greater efficiency during symbiosis (Terpolilli *et al*., 2008).

S. medicae inoculation results in the induction of the nodulation marker genes *ENOD11* and *NIN* (Figure 4.4). Both genes are very strongly transcribed in response to rhizobial inoculation, beginning 2-4 days after initial rhizobial inoculation and continue to exhibit elevated expression at the final 12 day timepoint at which point nodules are evident on rhizobia-treated roots. Expression of both of these genes within these timeframes is established in *S. meliloti-M. truncatula* symbiosis (Journet *et al*., 2001; Vernie *et al*., 2015). Given that these genes are known to be essential for nodulation, these expression patterns would be expected to be conserved between different compatible rhizobial species.

Although these expression patterns of *NIN* and *ENOD11* are broadly consistent with data from other rhizobia – *M. truncatula* interactions, there is evidence that both of these genes are actually expressed more rapidly within the context of the symbiosis. *ENOD11* expression has been detected by GUS staining of p*MtENOD11*-β-glucuronidase and RT-qPCR as early as two hours after Nod factor treatment in epidermal root hair cells (Charron *et al*., 2004; Journet *et al*., 2001). Expression of GUS under the *NIN* promoter also suggests that *NIN* is expressed exclusively within the epidermis 24 hours after Nod factor treatment (Vernie *et al*., 2015). Because the expression data presented here is from whole roots, it is possible that expression is not initially detectable when averaged out across the breadth of the whole root and only becomes significantly different later in the symbiosis when *NIN* and *ENOD11* are expressed in other layers of the root. This is consistent with the knowledge that both of these markers do become expressed in the cortex following infection thread formation (Journet *et al*., 2001; Yoro *et al*., 2014).

Putative defence markers do not respond to ralstonia, rhizobia or flg22: Surprisingly, expression of *MtFLS2* was not found to be increased by ralstonia (Figure 4.5 B) inoculation and only 2 DAI following rhizobia inoculation (Figure 4.5A). This increase seen 2 DAI could be an artefact of insufficient biological replication; although rhizobia have been demonstrated to induce defence

responses during nodulation as reviewed previously, this normally occurs at a much earlier point in the interaction.

It has been demonstrated that *R. solanacearum* does possess MAMPs that activate immune responses in *A. thaliana*. However, these responses are consistent between extracts from wild type *R. solanacearum* and *fliC* or *flhDC* mutants deficient in flagellin production. Furthermore, wild type bacteria are able to induce similar disease symptoms in both wild type and *fls2-101* mutants and a peptide based on the region of bacterial flagellin corresponding to flg22 does not elicit defence responses (Pfund *et al*., 2004). This suggests that *R. solanacearum* flagellin is not immunogenic and therefore if *MtFLS2* has the same role as its counterpart in *A. thaliana*, its lack of response to ralstonia could be explained.

From a review of the literature, it would be expected that rhizobia should also elicit defence responses, although it should be noted that *Mt*FLS2 is not induced by *S. meliloti* in the Chen *et al.* study either (Chen *et al*., 2017). However, the strength of rhizobial inoculations used in many of these studies are likely far higher than the plants would encounter naturally and thus the defence responses may not be physiologically relevant. For instance some studies have used an OD_{600} of 0.8 (Libault *et al*., 2010) or 1.0 (Lohar *et al*., 2006) when the OD_{600} of 0.05 used in these experiments is sufficient to result in nodule formation and nitrogen fixation (Figure 4.1G, Figure 4.3) and the reported induction of defence-related genes could be a consequence of overexposure to rhizobia. Expression of *LjWRKY33* and *LjMAPK3* have been reported following inoculation with OD_{600} 0.02 *M. loti* although MAPK phosphorylation in the same study required inoculation with OD_{600} 0.5 (Lopez-Gomez *et al*., 2012).

Although *FLS2* is induced during immune responses in *A. thaliana* (Gomez-Gomez and Boller, 2000), this induction is not usually potent when compared to other defence markers and the activity of the receptor is thought to be regulated predominantly at the posttranscriptional level (Withers and Dong, 2017). Therefore, the inclusion of *MtFLS2* in this study was based on the data suggesting that it was very strongly responsive to *P. syringae* in *M. truncatula* (Chen *et al*., 2017) rather than on the basis of the *A. thaliana* homolog being a robust marker of immune responses. Even if *R.*

solanacearum and *S. medicae* lack the required MAMPs to activate its expression, *MtFLS2* was not found to be induced by flg22 treatment either (Figure 4.6F). It may be possible that the response of *MtFLS2* to *P. syringae* inoculation is an effect associated specifically with this pathogen, and this could be explored in future work.

Experiments using flg22 also did not elicit any induction of the other putative defence markers tested (Figure 4.6). For example, a putative homolog of *PER4*, another flg22-responsive gene in *A. thaliana* (Boudsocq *et al*., 2010) had been identified for use here (but it was not ultimately tested due to unfavourable primer annealing properties). It had been used as a defence marker in a study (Rey *et al*., 2017) on the role of the GRAS protein RAD1 (Required for arbuscule development 1) moderating infection by the pathogen *Phytophthora palmivora*. This study also described another gene, Medtr1g075340, as a defence marker which was found not to respond to flg22 (figure 4.6A). Neither gene was shown to be up-regulated in response to *P. palmivora* however, and the basis on which the authors prescribe these genes as defence markers is unknown.

It should also be noted that although flg22 is a very potent inducer of some genes associated with innate immunity in *A. thaliana*, it is not certain that it elicits the same response in *M. truncatula*. However, flg22 has been convincingly demonstrated to induce similar pathways in the legumes *G. max* and *L. japonicus* which belong to the same sub-family (Faboideae) as *M. truncatula* (Liang *et al*., 2013; Lopez-Gomez *et al*., 2012). More specifically, flg22 treatment has been shown to induce the expression of the *DCL4* gene in *M. truncatula*. DCL4 belongs to the family of dicer-like proteins which are required for the generation of siRNAs and miRNAs (Tworak *et al*., 2016). *DCL4* was also shown to be induced in rhizobia-containing nodules although it is unclear if this is sufficient to suggest a link to innate immunity. Alternative approaches, such as measuring the production of ROS in the presence or absence of flg22 could be used to ask if the peptide is immunogenic in *M. truncatula*.

Interplay between circadian rhythms and plant defence: The RNAseq expression levels (Figure 4.8) and flg22 RT-qPCR expression levels (Figure

4.6) were calculated relative to 0 hour (immediately before treatment, see methods) samples instead of relative to mock treatments at the same timepoint. As treatment and sampling was synchronized such that the 0 hour timepoint occurred at dawn, this means that these samples are referenced against plants that are untreated but also collected at a different time of day. It is less likely that there will be significant developmental differences over such short periods of time but the possibility of circadian effects is certain.

In addition to the confounding effects of expression of genes regulated by the circadian clock in some of the datasets, which can be difficult to uncouple from treatment-specific effects, there is a further layer of complexity when studying plant-microbe interactions since these may themselves also be influenced by circadian rhythms. The presence or absence of light influences the effectiveness of defence responses, particularly those associated with the SA pathway (Roden and Ingle, 2009) which would likely include those induced by *R. solanacearum*. This has been elegantly demonstrated by experiments inoculating *A. thaliana* with *P. syringae* at different times of the day with plants showing increased susceptibility at night and greater resistance during the day (Bhardwaj *et al*., 2011). Overexpression of key clock genes *CCA1*/*LHY* in *A. thaliana* also leads to increased susceptibility to both *P. syringae* and *Hyaloperonospora arabidopsidis* (Zhang *et al*., 2013).

Whilst the 16 hour photoperiod used in these experiments ensures all samples were taken at either dawn or during the 'day,' the possibility that the response to ralstonia or rhizobia inoculation seen in the RNAseq experiments includes effects related to the circadian rhythm of the plant is high. This could be addressed in future work for by carrying out gene expression profiling using the mock-inoculated samples at the 3 and 12 HAI timepoints.

Insights into nitrogen metabolism in *M. truncatula*: Although these experiments were not designed to investigate N-responses, 90 genes that show altered expression following nitrogen provision can be identified by comparing the replete-/deplete-nitrogen T=0 samples which are essentially otherwise identical to mock-treated samples that have been grown on different N levels for 24 hours (see methods). 70 genes were found to be N-induced and 20 N-repressed.

The most strongly induced of these is Medtr2g069273, which is an NCR peptide and thus suggests a role for this peptide beyond regulating the bacterial population of nodules (as discussed earlier). There are also some genes putatively involved in nitrate and amino acid metabolism present. These include nitrate-reductases (Medtr5g059820 and Medtr3g073180), a putative ferredoxin-nitrate reductase (Medtr4g086020), a serine-glyoxylate aminotransferase (Medtr5g090070), a class I glutamine amidotransferase (Medtr7g082570), putative components of *NRT3.1* (Medtr4g104700 and Medtr4g104730) and an asparagine synthetase (Medtr3g464580). Amongst the down-regulated genes is an asparagine peptidase (Medtr1g018760) and a homolog of *AtNRT2.5* (Medtr8g069775). Genes annotated with the GO terms "iron ion binding," (P=0.019) "heme binding" (P=0.024) and "tetrapyrrole binding" (P=0.025) are significantly overrepresented amongst the nitrogen responsive genes. Although the most significant tetrapyrroles in plants are chlorophylls, these annotations are shared amongst the same group of 8 genes, suggesting that the tetrapyrrole in question is more likely to be a seroheme, a co-factor of nitrite and sulphite reducing enzymes (Vavilin and Vermaas, 2002).

The young age of the seedlings (3 days at time of sampling) and absence of timepoints closer to nitrogen provision used for RNAseq limits the detail in which N-responses can be assayed. Sequencing of additional mock samples from later timepoints and the inclusion of shorter term timepoints would allow N-responses to be interrogated throughout development of *M. truncatula* seedlings. Given the relationship between nitrogen nutrition and nodulation, a better understanding of the primary nitrate response and later reprogramming of nitrogen metabolism in legumes may also provide insight into nodulation (Lagunas *et al*., 2019).

Identification of defence-associated genes in *M. truncatula*: DE genes from the 12 HAI ralstonia samples would be expected to be enriched for defence-associated genes and this hypothesis is supported by the overrepresented GO term 'defence responses.' However, this is assigned to genes on the basis of inference from homology to other proteins, rather than experimental evidence and may include genes not related to defence whilst

omitting others, given the relatively poor understanding of defence signalling in *M. truncatula*. Therefore, in to attempt to identify novel defence-associated genes in this model, the DE genes from ralstonia 12 HAI sample were considered. To try to avoid the confounding effect from circadian changes in this dataset, DE genes intersecting with the 12 HAI rhizobia set were removed, leaving 505 genes (333 up-regulated, 172 down-regulated). Given that rhizobia treatment is associated with stimulation of immunity around this timescale (as discussed previously), removal of some defence-associated genes cannot be excluded by this approach. However, the GO term 'defence responses' was still associated with 27 genes in this list, contrasted to 31 in the original dataset, suggesting that the majority of defence-associated genes are preserved.

Two isoflavone synthases, *MtIFS1* (Medtr7g027960) and *MtIFS2* (Medtr7g028020), are both strongly induced by ralstonia 12 HAI. Isoflavones are secondary metabolites conserved throughout Fabaceae. Although, little is known about isoflavones specifically in *M. truncatula*, these compounds have been widely associated with direct antimicrobial activity and can act as phytoalexin precursors across the Fabaceae family (Goyal *et al*., 2012). Several predicted wall-associated kinases (WAKs) were present amongst these genes (Medtr1g110180, Medtr2g031530, Medtr3g088760). WAK proteins recognize pectin, both present in cell walls and in fragments and thus have roles in regulating cell growth and recognition of pathogen attack (Kohorn and Kohorn, 2012) and therefore their induction by ralstonia is plausible.

A calmodulin-dependent protein kinase (Medtr7g106710) with homology to the *AtCPK1* is present in this list. In *A. thaliana* this protein is closely associated with defence pathways; its expression is induced during immunity and its overexpression leads to induction of genes associated with the SA pathway whilst loss of function leads to increased pathogen susceptibility (Coca and San Segundo, 2010). Furthermore, *CPK1* is associated with ROS production and programmed-cell death during defence responses (Gao *et al*., 2013). Also present is a homolog of *AtSD2-5* (Medtr2g073250). This gene belongs to a family of proteins called *S*-domain receptor kinases that are associated with regulation of the self-incompatibility pathway in *A. thaliana* (Samuel *et al*., 2008). However, eFP browser data

shows this gene is potently induced by bacterial MAMPs including flg22 and harpin (Brunner *et al*., 2007) and may therefore also be involved in defence pathways.

In the 3 HAI ralstonia sample, one of the most up-regulated genes is a predicted pathogen-inducible alpha-dioxygenase (Medtr6g007770) which shares a high degree of conservation with the *A. thaliana* protein *DOX1* (dioxygenase 1). This protein is induced by various bacterial interactions and during the salicylic acid defence pathway (Blanco *et al*., 2005) and is understood to protect against oxidative stress during defence responses (De Leon *et al*., 2002).

Testing the expression of these defence-associated genes in response to a generic inducer of innate immunity, such as flg22, would ascertain if these genes are likely to act as common components of the *M. truncatula* defence response or if they are induced specifically by ralstonia. It would also be worthwhile in future work to test the expression of these in all of the rhizobia-inoculated plant samples by RT-qPCR to ask if they are associated with rhizobia-linked defence responses. The identification of defence markers in this species would be of broad benefit in understanding the role of immunity during nodulation and in studying *M. truncatula*-pathogen interactions.

Interpreting early responses to rhizobia: Overall, neither the 3 HAI and 12 HAI rhizobia datasets include many genes obviously associated with nodulation. By these timepoints, it is possible that a specific symbiotic response to rhizobia has not yet had a chance to form. However, expression of the marker *ENOD11* has been demonstrated in the epidermis within hours of rhizobial inoculation as previously discussed. This lack of finding of strong gene expression changes could again be attributable to this dataset being generated from whole root samples because mixed patterns of induction and repression/low expression obscure cell type specific effects, such as found with early responses to rhizobia in the epidermis. The RNAseq data is in line with the finding of an absence of expression of symbiosis markers in earlier timepoints from the RT-qPCR data (Figure 4.4).

A predicted zinc transporter with high homology with the *A. thaliana* zinc transporter *ZIP10* (ZRT/IRT-like protein 10) is activated 3 HAI with rhizobia.

This gene (Medtr4g083570) has been characterized as zinc permease *ZIP6* in *M. truncatula* which is expressed exclusively in roots and nodules. Zinc is required as a co-factor for rhizobial nitrogenase and RNA interference of this gene results ineffective nitrogen fixation, suggesting it is involved in supplying zinc to fixing rhizobia (Abreu *et al*., 2017). However, this role is not consistent with expression at such an early timepoint. It is possible that this protein has additional roles in nodulation, possibly in root hair curling, given that *A. thaliana ZIP10* is a regulator of root hair development (Di Cristina *et al*., 1996). It would be interesting to follow up on this finding.

A higher than expected proportion of the genes in the 12 HAI rhizobia dataset were annotated with the GO term 'lipid transport' due to the presence of 6 putative lipid transfer proteins. Other lipid transfer proteins have been demonstrated to be involved in nodulation. For instance, the *M. truncatula* protein *N5* has been demonstrated to be a positive regulator of nodulation; silencing of this gene leads to a significant (but not total) reduction in nodule number and its overexpression leads to hypernodulation. The gene is epidermis-specific and it is induced within days of *S. meliloti* inoculation (Pii *et al*., 2009; Pii *et al*., 2012). It is likely that other lipid transfer proteins have roles in nodulation during the early stages of the symbiosis, given that significant cellular reorganization is required for infection thread formation and preparation for division of cortical cells.

Although the GO term 'lipid transport' is not overrepresented amongst the 3 HAI rhizobia samples, some proteins related to this process are conserved between this and the 12 HAI dataset. These include a lipid transfer protein (Medtr1g012700) and a GDSL-like lipase (Medtr5g074270) which are the first and third most strongly induced genes in this dataset once circadian genes are excluded, reinforcing the importance of lipid remodelling during early nodulation. Genes likely to be associated with carbohydrate metabolism are also present at this timepoint; these include proteins with high homology to *A. thaliana* galactinol synthases *GolS1* (Medtr2g084340) and *GolS2* (Medtr7g111850) and a predicted beta-amylase (Medtr5g013620). *A. thaliana* galactinol synthases are associated with heat shock responses (Panikulangara *et al*., 2004) and two predicted heat shock proteins

(Medtr4g130540 and Medtr1g015165) are also up-regulated, however the functional significance of this is unclear.

Interplay between defence and symbiosis in *M. truncatula*: Direct comparison between the rhizobia and ralstonia datasets is very difficult because the plants sampled for these were grown in vastly different N-conditions and due to the confounding circadian effects, as discussed previously. Despite these caveats, DE genes were calculated between the rhizobia and ralstonia samples at the 3 HAI and 12 HAI timepoints. Genes associated with differences in N-level as calculated between the reference samples were removed from these datasets. The circadian-associated genes were not removed as comparisons were made at the same timepoint and any differences in their expression may be relevant to interpreting the data (and include inoculation/circadian interaction effects). However, none of the 5 genes conserved between rhizobia and ralstonia 3 HAI were DE between the two treatments and only 7 of the 117 genes conserved between the two 12 HAI datasets were DE *between* the treatments. This left 205 genes DE between rhizobia and ralstonia at 3 HAI and 325 at 12 HAI with an overlap of 27 DE genes between the two sets.

GO term analysis of these genes did show the presence of some overrepresented processes amongst the 3 HAI comparison. This included 7 genes associated with 'lipid transport' (P=3.9e-03) which are all repressed by ralstonia relative to rhizobia. At 12 HAI, there were no processes significantly enriched amongst the DE genes but many of the most strongly induced or repressed genes are hypothetical proteins or poorly functionally annotated with no well characterized homologs in other plant species.

One gene that in repressed by ralstonia relative to rhizobia is *LYK2* (Medtr5g086330) which is an ortholog of the Nod factor receptor *LYK3* (De Mita *et al*., 2014) with which is possibly functionally redundant (Buendia *et al*., 2018). Although this gene is DE between these two treatments but not between each treatment and its respective mock, it is possible that this effect may be much more pronounced at different timepoints or if cell type specific expression data was available. Nevertheless, the possibility of a receptor

associated with symbiosis being repressed during pathogenesis is very interesting in the context of legume-microbe interactions.

Although only being DE in the 12 HAI ralstonia dataset, *PT4* (Medtr1g028600) is particularly noteworthy. This gene encodes a phosphate transporter (Harrison *et al*., 2002) which is essential for phosphate uptake in *M. truncatula* plants during mycorrhizal symbiosis and is considered a marker gene for this interaction (Javot *et al*., 2007). Activation of this gene in response to ralstonia inoculation is unexpected, especially given that all plants were grown in replete phosphate and the possible P-N interplay could be explored.

Conclusions: The response of *M. truncatula* seedlings to the pathogen *R. solanacearum* and symbiont *S. medicae* were compared over time. *R. solanacearum* induced stunting of plant growth within 2 weeks of inoculation and led to wilting and death of infected plants within 4 weeks (Figure 4.2). Nodules were formed on roots treated with rhizobia within 9 days of inoculation (Figure 4.1G).

RT-qPCR analysis showed the induction of nodulation marker genes *NIN* and *ENOD11* at later timepoints in these experiments (Figure 4.4) in response to *S. medicae*, confirming these genes are induced in a similar fashion as with other rhizobial strains. However, the *M. truncatula* homolog of *FLS2* was not induced by rhizobia, ralstonia or flg22 and other putative defence markers were also not flg22-responsive within the timescales of these experiments.

RNAseq was used to identify genes associated with inoculation with rhizobia and ralstonia 3 and 12 hours after inoculation. Functional analysis of DE genes was hindered by the lack of information available on the activity of many genes in the *M. truncatula* genome. However, genes were identified that may function as novel markers of defence responses in *M. truncatula* that merit further study (e.g. *MtIFS1*/*MtIFS2*, *MtCPK1*). Some genes associated with nitrogen availability and rhizobia inoculation were also identified.

Comparisons between the rhizobia and ralstonia datasets were complicated by the influence of circadian effects in addition to the bacterial inoculations. Nevertheless, the possibility of genes that are differentially regulated during symbiosis and defence and may therefore be regulators of

permitting symbiosis to occur in place of defence responses normally associated with plant-microbe interactions is worthwhile to study further.

Chapter 5: General discussion:

Insights into how the environment shapes plant root architecture: Plant root systems exhibit a remarkable level of plasticity as they adapt to their environment. Two specific examples of such alterations underpinning gene expression changes involved in their regulation were investigated in lateral root formation and nodule formation, both of which adaptations are subject to regulation from both abiotic and biotic inputs.

To identify genes involved in the regulation of lateral root emergence in *A. thaliana*, FACS was employed over a timecourse experiment (Chapter 3: Regulation of gene expression by nitrogen and rhizobia underpinning lateral root emergence in *Arabidopsis thaliana*). The responses to treatments associated with lateral root formation in *A. thaliana* in nitrate provision and *S. meliloti* inoculation were examined in the pericycle and cortex. These responses were found to be highly cell type specific with little conservation between treatment-responsive genes in cells of the cortex and pericycle. Furthermore, trends in the regulation of these responses such as possible signal transduction of the nitrate response throughout the root and apparent repression of genes associated with the primary nitrate response in the pericycle were unearthed that are masked at the whole organ or organism level (Walker *et al*., 2017).

Lateral root emergence is known to depend on the integration of multiple positive and negative stimuli (Benkova and Bielach, 2010) that induce distinct transcriptional reprogramming in different cell types (Gifford *et al*., 2008). These experiments reinforce the importance of cell type specificity during complex developmental responses. Obtaining data for the cell types of the root absent from this dataset would enable the regulation of lateral root emergence to be examined at an even higher temporal and spatial resolution than has ever been achieved previously, possibly identifying further novel and biologically relevant mechanisms.

Next, to obtain insight into how the competing demands of protecting against pathogen attack and forming beneficial interactions with symbionts are balanced in *M. truncatula*, gene expression was assayed following inoculation with the symbiont *S. medicae* or the pathogen *R. solanacearum* (Chapter 4:

molecular changes underpinning 'friend vs. foe' recognition in legume roots). Due to *M. truncatula* rarely being used as a model for plant-pathogen interactions and the consequential lack of knowledge of genes involved in the regulation of defence pathways, interactions between symbiosis and defence were difficult to confirm. An approach based on identifying homologs of defence-associated genes in the better studied *A. thaliana* model did not conclusively identify suitable defence markers, although some candidates were identified by RNAseq analysis of *R. solanacearum* inoculated roots.

Despite progress in understanding defence responses in legumes, more work is needed in this area before the question of how defence is regulated during nodulation can be fully answered. Additionally, the approach used in these experiments to study regulation at the whole root level may not identify all biologically important changes in gene expression, especially if these changes are restricted to single cell types, as would be expected based on the results of Chapter 3. Nevertheless, the data obtained from these experiments contribute to understanding pathogen responses in *M. truncatula*.

Future directions of this research:

Characterizing rhizobia responsive genes in non-legumes: In this work, RNAseq analysis was used to identify rhizobia responsive genes during early interactions between *S. medicae* and *M. truncatula*. Very few of these genes are so far well characterized, but some of these would make interesting candidates for further study. In particular, the proteins annotated as being involved in lipid transport could plausibly have important roles in nodulation, given the importance of other lipid regulatory proteins in the symbiosis (Pii *et al*., 2012).

Additionally, having already characterized the response of the non-legume *A. thaliana* to another rhizobial species, *S. meliloti*, it may also be interesting to compare these datasets. Although the *A. thaliana* dataset is cell type-specific it does feature samples taken at similar timepoints after inoculation compared to the *M. truncatula* RNAseq data. Although *A. thaliana* is not able to form nodules, the data presented in Chapter 3 demonstrate both

transcriptional and developmental responses to rhizobia that would be interesting to examine.

Additionally, other studies have demonstrated the suppression of defence responses in *A. thaliana* by rhizobial Nod factors, a response that is conserved in legumes. This is made especially interesting by the fact that this response has been demonstrated to depend on the function of the receptor gene *LYK3* which shares high conservation with some legume Nod factor receptors (Liang *et al*., 2013). This raises the possibility that non-legumes are able to perceive and respond specifically to rhizobia to at least some extent. It would therefore be interesting to look at the expression patterns for homologs of genes associated with rhizobial inoculation in legumes in *A. thaliana* and other non-legumes. This would provide a deeper understanding of the extent to which non-legumes can interact with rhizobia which may contribute towards the goal of ultimately transferring nodulation to non-legume species.

Towards identifying defence markers in *M. truncatula*: Attempts to investigate defence responses in *M. truncatula* have been made difficult by the lack of knowledge of genes that may act as defence markers in this organism. Despite poor functional annotation of many genes regulated in response to the pathogen *R. solanacearum* and therefore possibly related to defence, RNAseq analysis has identified some genes that show differential expression which, based on homology to *A. thaliana* proteins, may function in defence responses. These include possible isoflavone synthases *MtIFS1*/*MtIFS2*, wall-associated kinases and a calmodulin-dependent kinase *MtCPK1*.

RT-qPCR could be used to test their expression to validate their differential expression in response to *R. solanacearum*. This could also be carried out in more samples to see how these genes behave in the more immediate and also long-term interactions with the pathogen. Furthermore, their expression in response to *S. medicae* and flg22 could also be investigated to gather data at the transcriptional level of the regulation of defence responses by rhizobia. This would also help clarify if the induction of these genes is related to a general defence response or is specific to the *M. truncatula*/*R. solanacearum* interaction.

Another approach to identify *M. truncatula* defence markers could involve using meta analysis of published datasets characterizing the transcriptomic response of the plant to a range of different pathogens. Comparison to older datasets based on outdated annotations of the *M. truncatula* genome is difficult and studies employing microarrays may not use probes with complete coverage of the *M. truncatula* transcriptome. However, there are numerous suitable datasets, including the RNAseq dataset presented here, that could be used for such an analysis. For instance, transcriptome studies have been published detailing the response of *M. truncatula* to the necrotrophic fungal pathogens *Macrophomina phaseolina* (Mah *et al*., 2012), *Fusarium oxysporum* (Thatcher *et al*., 2016) and *Rhizoctonia solani* (Liu *et al*., 2017). In the latter two cases, the interaction with *M. truncatula* A17 is resistant thus meaning a robust, protective defence response is taking place.

Application of fluorescence activated cell sorting to non-Arabidopsis species: No published studies utilizing FACS in plants outside of the model organism *A. thaliana* were known to exist at the time of writing. Other plant models, such as *M. truncatula*, are often less amenable to stable transformation and more commonly exhibit silencing of transgenes between generations (Rajeevkumar *et al*., 2015), thus making the generation of stable marker lines more difficult. Furthermore, most plant species produce fewer seeds, cannot be grown at such high density and may have cell wall compositions that are more resistant to enzymatic digestion relative to *A. thaliana*; all of these increase the difficulty in obtaining a sufficiently high protoplast yield for downstream analysis, particularly when rare target cell types are concerned.

Although *A. thaliana* is an excellent model for study of many root developmental processes, it is not suitable for the study of root symbiotic interactions with microorganisms such as mycorrhizae and rhizobia as is it capable of neither mycorrhization or nodulation. Despite an innovative approach to studying nodulation in *M. truncatula* root hairs (Breakspear *et al*., 2014), little is known about how individual root cell types behave during symbiosis. In order to gain maximum insight into how these processes operate

at the cell type-specific level, it is therefore necessary to utilize cell type-specific approaches in model plants aside from *A. thaliana*.

Nodulation represents a process that requires different cell types to commit to vastly different but co-ordinated developmental programmes. Therefore, it would be ideal to understand transcriptional reprogramming for each cell type. Indeed, it is understood that many of the genes involved in the activation of nodulation do show different behaviours between the cortex and the epidermis (Hayashi *et al*., 2014). That such an approach would be useful is illustrated by the behaviour of nodulation marker genes in these experiments (Figure 4.4). As discussed previously, these genes are found to be activated days after rhizobial inoculation despite other studies demonstrating their expression, specifically in the epidermis, within hours. Therefore, it is likely that many other biologically relevant mechanisms are at play in a cell type-specific manner early on during *M. truncatula*-microbe interactions, including at the timepoints used in the RNAseq experiments, that are not detected due to expression profile being measured from the entirety of the root.

If the aforementioned difficulties could be overcome, a FACS-based approach to studying nodulation would provide much more detail and clarity on how nodule development or defence responses are co-ordinated or mitigated. Cell types of particular interest would be the epidermis (as the outermost layer which first encounters rhizobia in the soil), the cortex (the site of nodule emergence) and the pericycle (the site of lateral root emergence which is theorized the be evolutionarily related to nodulation (Bright *et al*., 2005)).

Broader impact of findings:

Nitrogen is the most growth-limiting compound for plants and typically has to be provided to plants in the form of fertilizer in intensive agricultural systems (Vance, 2001). In Western countries, the majority of this fertilizer is produced industrially which is expensive and energy intensive. The production of ammonia for fertilizer for agricultural use via the Haber-Bosch process accounts for an estimated 1% of total global energy use. Further environmental damage is also associated with runoff of fertilizer into waterways during its application to crops (Conley *et al*., 2009). With the global

population expected to exceed 9 billion people by 2050, providing sufficient food for this many people without allocating even greater amounts of land towards agriculture remains a significant challenge.

For this reason, genetically modifying non-legumes to enable them to form symbiosis with nitrogen-fixing bacteria is an attractive option if GMO use is allowed in the future. In particular, transferring nodulation to cereal crops is a key objective of research into the symbiosis as fertilizer usage could be significantly reduced with self-nitrifying crops. Efforts are already underway to achieve this and these efforts tend to focus on the transfer of the legume nodulation machinery to other species (Charpentier and Oldroyd, 2010), but do not consider concomitant responses to pathogens in plants. Understanding how legumes are able to partition defence and symbiosis responses during nodulation may also be necessary step if the ability the form to nodules is to be transferred to non-legumes.

References:

Abreu, I., Saez, A., Castro-Rodriguez, R., Escudero, V., Rodriguez-Haas, B., Senovilla, M., Larue, C., Grolimund, D., Tejada-Jimenez, M., Imperial, J*., et al.* (2017). *Medicago truncatula* Zinc-Iron Permease6 provides zinc to rhizobia-infected nodule cells. Plant Cell Environ *40*, 2706-2719.

Acosta-Jurado, S., Rodriguez-Navarro, D.N., Kawaharada, Y., Perea, J.F., Gil-Serrano, A., Jin, H., An, Q., Rodriguez-Carvajal, M.A., Andersen, S.U., Sandal, N*., et al.* (2016). *Sinorhizobium fredii* HH103 Invades Lotus burttii by Crack Entry in a Nod Factor-and Surface Polysaccharide-Dependent Manner. Mol Plant Microbe Interact *29*, 925-937.

Albrecht, C., Geurts, R., and Bisseling, T. (1999). Legume nodulation and mycorrhizae formation; two extremes in host specificity meet. EMBO J *18*, 281-288.

Aldon, D., Brito, B., Boucher, C., and Genin, S. (2000). A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. EMBO J *19*, 2304-2314.

Almagro, A., Lin, S.H., and Tsay, Y.F. (2008). Characterization of the Arabidopsis nitrate transporter *NRT1.6* reveals a role of nitrate in early embryo development. Plant Cell *20*, 3289-3299.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Alvarez, J.M., Riveras, E., Vidal, E.A., Gras, D.E., Contreras-Lopez, O., Tamayo, K.P., Aceituno, F., Gomez, I., Ruffel, S., Lejay, L*., et al.* (2014). Systems approach identifies *TGA1* and *TGA4* transcription factors as important regulatory components of the nitrate response of *Arabidopsis thaliana* roots. Plant J *80*, 1-13.

Amor, B.B., Shaw, S.L., Oldroyd, G.E., Maillet, F., Penmetsa, R.V., Cook, D., Long, S.R., Denarie, J., and Gough, C. (2003). The *NFP* locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. Plant J *34*, 495-506.

Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N.H., Zhu, S., Qiu, J.L., Micheelsen, P., Rocher, A., Petersen, M*., et al.* (2005). The MAP kinase substrate *MKS1* is a regulator of plant defense responses. EMBO J *24*, 2579-2589.

Angelini, C., Cutillo, L., De Canditiis, D., Mutarelli, M., and Pensky, M. (2008). BATS: a Bayesian user-friendly software for analyzing time series microarray experiments. BMC Bioinformatics *9*, 415.

Angot, A., Peeters, N., Lechner, E., Vailleau, F., Baud, C., Gentzbittel, L., Sartorel, E., Genschik, P., Boucher, C., and Genin, S. (2006). *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. Proc Natl Acad Sci U S A *103*, 14620-14625.

Axtell, M.J., and Staskawicz, B.J. (2003). Initiation of *RPS2*-specified disease resistance in Arabidopsis is coupled to the *AvrRpt2*-directed elimination of *RIN4*. Cell *112*, 369-377.

Babe, A., Lavigne, T., Severin, J.P., Nagel, K.A., Walter, A., Chaumont, F., Batoko, H., Beeckman, T., and Draye, X. (2012). Repression of early lateral root initiation events by transient water deficit in barley and maize. Philos Trans R Soc Lond B Biol Sci *367*, 1534-1541.

Banda, J., Bellande, K., von Wangenheim, D., Goh, T., Guyomarc'h, S., Laplaze, L., and Bennett, M.J. (2019). Lateral Root Formation in Arabidopsis: A Well-Ordered LRexit. Trends Plant Sci *24*, 826-839.

Bao, Y., Aggarwal, P., Robbins, N.E., 2nd, Sturrock, C.J., Thompson, M.C., Tan, H.Q., Tham, C., Duan, L., Rodriguez, P.L., Vernoux, T*., et al.* (2014). Plant roots use a patterning mechanism to position lateral root branches toward available water. Proc Natl Acad Sci U S A *111*, 9319-9324.

Barbulova, A., Rogato, A., D'Apuzzo, E., Omrane, S., and Chiurazzi, M. (2007). Differential effects of combined N sources on early steps of the Nod factor-dependent transduction pathway in *Lotus japonicus*. Mol Plant Microbe Interact *20*, 994-1003.

Bargmann, B.O., Marshall-Colon, A., Efroni, I., Ruffel, S., Birnbaum, K.D., Coruzzi, G.M., and Krouk, G. (2013a). TARGET: a transient transformation system for genome-wide transcription factor target discovery. Mol Plant *6*, 978-980.

Bargmann, B.O., Vanneste, S., Krouk, G., Nawy, T., Efroni, I., Shani, E., Choe, G., Friml, J., Bergmann, D.C., Estelle, M*., et al.* (2013b). A map of cell type-specific auxin responses. Mol Syst Biol *9*, 688.

Bari, R., and Jones, J.D. (2009). Role of plant hormones in plant defence responses. Plant Mol Biol *69*, 473-488.

Bauer, W.D., and Caetano-Anollés, G. (1990). Chemotaxis, induced gene expression and competitiveness in the rhizosphere. Plant and Soil *129*, 45-52.

Bednarek, P., Pislewska-Bednarek, M., Ver Loren van Themaat, E., Maddula, R.K., Svatos, A., and Schulze-Lefert, P. (2011). Conservation and clade-specific diversification of pathogen-inducible tryptophan and indole glucosinolate metabolism in *Arabidopsis thaliana* relatives. New Phytol *192*, 713-726.

Bellini, C., Pacurar, D.I., and Perrone, I. (2014). Adventitious roots and lateral roots: similarities and differences. Annu Rev Plant Biol *65*, 639-666.

Ben, C., Debelle, F., Berges, H., Bellec, A., Jardinaud, M.F., Anson, P., Huguet, T., Gentzbittel, L., and Vailleau, F. (2013). *MtQRRS1*, an R-locus required for *Medicago truncatula* quantitative resistance to *Ralstonia solanacearum*. New Phytol *199*, 758-772.

Benkova, E., and Bielach, A. (2010). Lateral root organogenesis - from cell to organ. Curr Opin Plant Biol *13*, 677-683.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E. (2015). The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis *53*, 474-485.

Berendsen, R., L., Pieterse, C.M.J., and Bakker, P.A.H.M. (2012). The rhizosphere microbiome and plant health. Trends in Plant Science *17*, 478.

Berrabah, F., Ratet, P., and Gourion, B. (2015). Multiple steps control immunity during the intracellular accommodation of rhizobia. Journal of Experimental Botany *66*, 1977-1985.

Bethke, G., Unthan, T., Uhrig, J.F., Poschl, Y., Gust, A.A., Scheel, D., and Lee, J. (2009). Flg22 regulates the release of an ethylene response factor substrate from MAP kinase 6 in *Arabidopsis thaliana* via ethylene signaling. Proc Natl Acad Sci U S A *106*, 8067-8072.

Bhardwaj, V., Meier, S., Petersen, L.N., Ingle, R.A., and Roden, L.C. (2011). Defence responses of *Arabidopsis thaliana* to infection by *Pseudomonas syringae* are regulated by the circadian clock. PLoS One *6*, e26968.

Birnbaum, K., Jung, J.W., Wang, J.Y., Lambert, G.M., Hirst, J.A., Galbraith, D.W., and Benfey, P.N. (2005). Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. Nat Methods *2*, 615-619.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A gene expression map of the Arabidopsis root. Science *302*, 1956-1960.

Blanco, F., Garreton, V., Frey, N., Dominguez, C., Perez-Acle, T., Van der Straeten, D., Jordana, X., and Holuigue, L. (2005). Identification of *NPR1*-dependent and independent genes early induced by salicylic acid treatment in Arabidopsis. Plant Mol Biol *59*, 927-944.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Boller, T., and Felix, G. (2009). A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. The Annual Review of Plant Biology *60*, 379-406.

Bolwell, G.P. (1999). Role of active oxygen species and NO in plant defence responses. Curr Opin Plant Biol *2*, 287-294.

Boucher, C.A., Barberis, P.A., Trigalet, A., and Demery, D.A. (1985). Transposon Mutagenesis of *Pseudomonas solanacearum*: Isolation of Tn5-Induced Avirulent Mutants. Journal of General Microbiology *131*, 2249-2457.

Boudsocq, M., Willmann, M.R., McCormack, M., Lee, H., Shan, L., He, P., Bush, J., Cheng, S.H., and Sheen, J. (2010). Differential innate immune signalling via Ca(2+) sensor protein kinases. Nature *464*, 418-422.

Brady, S.M., Orlando, D.A., Lee, J.Y., Wang, J.Y., Koch, J., Dinneny, J.R., Mace, D., Ohler, U., and Benfey, P.N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. Science *318*, 801-806.

Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., Morieri, G., Mysore, K.S., Wen, J., Oldroyd, G.E.*, et al.* (2014). The root hair "infectome" of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for Auxin signaling in rhizobial infection. Plant Cell *26*, 4680-4701.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y.S., Penfold, C.A., Jenkins, D.*, et al.* (2011). High-resolution temporal profiling of transcripts during Arabidopsis leaf senescence reveals a distinct chronology of processes and regulation. Plant Cell *23*, 873-894.

Brewin, N.J. (2004). Plant Cell Wall Remodelling in the Rhizobium–Legume Symbiosis. Critical Reviews in Plant Sciences *23*, 293-316.

Bright, L.J., Liang, Y., Mitchell, D.M., and Harris, J.M. (2005). The LATD gene of *Medicago truncatula* is required for both nodule and root development. Mol Plant Microbe Interact *18*, 521-532.

Brown, D.G., Swanson, J.K., and Allen, C. (2007). Two host-induced R*alstonia solanacearum* genes, *acrA* and *dinF*, encode multidrug efflux pumps and contribute to bacterial wilt virulence. Appl Environ Microbiol *73*, 2777-2786.

Brunner, F., Nürnberger, T., Townsend, H., Emmerson, Z., and Schildknecht, B. (2007). AtGenExpress: Response to bacterial-(LPS, HrpZ, Flg22) and oomycete-(NPP1) derived elicitors.

Buendia, L., Girardin, A., Wang, T., Cottret, L., and Lefebvre, B. (2018). LysM Receptor-Like Kinase and LysM Receptor-Like Protein Families: An Update on Phylogeny and Functional Characterization. Front Plant Sci *9*, 1531.

Canales, J., Moyano, T.C., Villarroel, E., and Gutierrez, R.A. (2014). Systems analysis of transcriptome data provides new hypotheses about Arabidopsis root response to nitrate treatments. Front Plant Sci *5*, 22.

Cao, Y., Halane, M.K., Gassmann, W., and Stacey, G. (2017). The Role of Plant Innate Immunity in the Legume-Rhizobium Symbiosis. Annu Rev Plant Biol *68*, 535-561.

Cao, Y., Liang, Y., Tanaka, K., Nguyen, C.T., Jedrzejczak, R.P., Joachimiak, A., and Stacey, G. (2014). The kinase *LYK5* is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase *CERK1*. Elife *3*.

Cárdenas, L., Martínez, A., Sánchez, F., and Quinto, C. (2008). Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). The Plant Journal *56*, 802-813.

Carter, A.D., Bonyadi, R., and Gifford, M.L. (2013). The use of fluorescence-activated cell sorting in studying plant development and environmental responses. Int J Dev Biol *57*, 545-552.

Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. Bioinformatics *26*, 2363-2367.

Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P., Sandberg, G., and Bennett, M.J. (2003). Dissecting Arabidopsis lateral root development. Trends Plant Sci *8*, 165-171.

Castaings, L., Camargo, A., Pocholle, D., Gaudon, V., Texier, Y., Boutet-Mercey, S., Taconnat, L., Renou, J.P., Daniel-Vedele, F., Fernandez, E.*, et al.* (2009). The nodule inception-like protein 7 modulates nitrate sensing and metabolism in Arabidopsis. Plant J *57*, 426-435.

Cerri, M.R., Frances, L., Laloum, T., Auriac, M.C., Niebel, A., Oldroyd, G.E., Barker, D.G., Fournier, J., and de Carvalho-Niebel, F. (2012). *Medicago truncatula* ERN transcription factors: regulatory interplay with *NSP1*/*NSP2* GRAS factors and expression dynamics throughout rhizobial infection. Plant Physiol *160*, 2155-2172.

Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E., and Parniske, M. (2008). *Lotus japonicus CASTOR* and *POLLUX* are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. Plant Cell *20*, 3467-3479.

Charpentier, M., and Oldroyd, G. (2010). How close are we to nitrogen-fixing cereals? Curr Opin Plant Biol *13*, 556-564.

Charpentier, M., and Oldroyd, G.E. (2013). Nuclear calcium signaling in plants. Plant Physiol *163*, 496-503.

Charron, D., Pingret, J.L., Chabaud, M., Journet, E.P., and Barker, D.G. (2004). Pharmacological evidence that multiple phospholipid signaling pathways link Rhizobium nodulation factor perception in *Medicago truncatula* root hairs to intracellular responses, including Ca2+ spiking and specific ENOD gene expression. Plant Physiol *136*, 3582-3593.

Chen, C.Z., Lv, X.F., Li, J.Y., Yi, H.Y., and Gong, J.M. (2012). Arabidopsis *NRT1.5* is another essential component in the regulation of nitrate reallocation and stress tolerance. Plant Physiol *159*, 1582-1590.

Chen, T., Duan, L., Zhou, B., Yu, H., Zhu, H., Cao, Y., and Zhang, Z. (2017). Interplay of Pathogen-Induced Defense Responses and Symbiotic Establishment in *Medicago truncatula*. Front Microbiol *8*, 973.

Cheng, H.P., and Walker, G.C. (1998). Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol *180*, 5183-5191.

Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The Arabidopsis receptor kinase *FLS2* binds flg22 and determines the specificity of flagellin perception. Plant Cell *18*, 465-476.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nurnberger, T., Jones, J.D., Felix, G., and Boller, T. (2007). A flagellin-induced complex of the receptor *FLS2* and *BAK1* initiates plant defence. Nature *448*, 497-500.

Cho, M.J., and Harper, J.E. (1991). Effect of inoculation and nitrogen on isoflavonoid concentration in wild-type and nodulation-mutant soybean roots. Plant Physiol *95*, 435-442.

Clavijo, F., Diedhiou, I., Vaissayre, V., Brottier, L., Acolatse, J., Moukouanga, D., Crabos, A., Auguy, F., Franche, C., Gherbi, H.*, et al.* (2015). The Casuarina *NIN* gene is transcriptionally activated throughout Frankia root infection as well as in response to bacterial diffusible signals. New Phytol *208*, 887-903.

Clua, J., Roda, C., Zanetti, M.E., and Blanco, F.A. (2018). Compatibility between Legumes and Rhizobia for the Establishment of a Successful Nitrogen-Fixing Symbiosis. Genes (Basel) *9*.

Cnops, G., Neyt, P., Raes, J., Petrarulo, M., Nelissen, H., Malenica, N., Luschnig, C., Tietz, O., Ditengou, F., Palme, K.*, et al.* (2006). The *TORNADO1* and *TORNADO2* genes function in several patterning processes during early leaf development in Arabidopsis thaliana. Plant Cell *18*, 852-866.

Coca, M., and San Segundo, B. (2010). *AtCPK1* calcium-dependent protein kinase mediates pathogen resistance in Arabidopsis. Plant J *63*, 526-540.

Conley, D.J., Paerl, H.W., Howarth, R.W., Boesch, D.F., Seitzinger, S.P., Havens, K.E., Lancelot, C., and Likens, G.E. (2009). Ecology. Controlling eutrophication: nitrogen and phosphorus. Science *323*, 1014-1015.

Cooper, J.E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. J Appl Microbiol *103*, 1355-1365.

Corcilius, L., Hastwell, A.H., Zhang, M., Williams, J., Mackay, J.P., Gresshoff, P.M., Ferguson, B.J., and Payne, R.J. (2017). Arabinosylation Modulates the Growth-Regulating Activity of the Peptide Hormone CLE40a from Soybean. Cell Chem Biol *24*, 1347-1355 e1347.

Crawford, N.M. (1995). Nitrate: nutrient and signal for plant growth. Plant Cell *7*, 859-868.

Cruz, C., Bio, A.F., Dominguez-Valdivia, M.D., Aparicio-Tejo, P.M., Lamsfus, C., and Martins-Loucao, M.A. (2006). How does glutamine synthetase activity determine plant tolerance to ammonium? Planta *223*, 1068-1080.

Cui, H., Tsuda, K., and Parker, J.E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. Annu Rev Plant Biol *66*, 487-511.

De Leon, I.P., Sanz, A., Hamberg, M., and Castresana, C. (2002). Involvement of the Arabidopsis alpha-*DOX1* fatty acid dioxygenase in protection against oxidative stress and cell death. Plant J *29*, 61-62.

De Mita, S., Streng, A., Bisseling, T., and Geurts, R. (2014). Evolution of a symbiotic receptor through gene duplications in the legume-rhizobium mutualism. New Phytol *201*, 961-972.

De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S*., et al.* (2010). A novel aux/*IAA28* signaling cascade activates *GATA23*-dependent specification of lateral root founder cell identity. Curr Biol *20*, 1697-1706.

De Smet, I., Signora, L., Beeckman, T., Inze, D., Foyer, C.H., and Zhang, H. (2003). An abscisic acid-sensitive checkpoint in lateral root development of Arabidopsis. Plant J *33*, 543-555.

De Smet, I., Vanneste, S., Inze, D., and Beeckman, T. (2006). Lateral root initiation or the birth of a new meristem. Plant Mol Biol *60*, 871-887.

De Smet, I., Vassileva, V., De Rybel, B., Levesque, M.P., Grunewald, W., Van Damme, D., Van Noorden, G., Naudts, M., Van Isterdael, G., De Clercq, R*., et al.* (2008). Receptor-like kinase *ACR4* restricts formative cell divisions in the Arabidopsis root. Science *322*, 594-597.

Deak, K.I., and Malamy, J. (2005). Osmotic regulation of root system architecture. Plant J *43*, 17-28.

Deakin, W.J., and Broughton, W.J. (2009). Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. Nat Rev Microbiol *7*, 312-320.

Deal, R.B., and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana. Nat Protoc *6*, 56-68.

Debelle, F., Maillet, F., Vasse, J., Rosenberg, C., de Billy, F., Truchet, G., Denarie, J., and Ausubel, F.M. (1988). Interference between *Rhizobium meliloti* and *Rhizobium trifolii* nodulation genes: genetic basis of *R. meliloti dominance*. J Bacteriol *170*, 5718-5727.

Dechorgnat, J., Nguyen, C.T., Armengaud, P., Jossier, M., Diatloff, E., Filleur, S., and Daniel-Vedele, F. (2011). From the soil to the seeds: the long journey of nitrate in plants. J Exp Bot *62*, 1349-1359.

Denison, R.F., and Kiers, E.T. (2004). Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. FEMS Microbiol Lett *237*, 187-193.

Desikan, R., Hancock, J.T., Ichimura, K., Shinozaki, K., and Neill, S.J. (2001). Harpin induces activation of the Arabidopsis mitogen-activated protein kinases *AtMPK4* and *AtMPK6.* Plant Physiol *126*, 1579-1587.

Di Cristina, M., Sessa, G., Dolan, L., Linstead, P., Baima, S., Ruberti, I., and Morelli, G. (1996). The Arabidopsis *Athb-10* (*GLABRA2*) is an HD-Zip protein required for regulation of root hair development. Plant J *10*, 393-402.

Di Laurenzio, L., Wysocka-Diller, J., Malamy, J.E., Pysh, L., Helariutta, Y., Freshour, G., Hahn, M.G., Feldmann, K.A., and Benfey, P.N. (1996). The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell *86*, 423-433.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Downie, J.A. (2014). Legume nodulation. Curr Biol *24*, R184-190.

Droillard, M.J., Boudsocq, M., Barbier-Brygoo, H., and Lauriere, C. (2004). Involvement of *MPK4* in osmotic stress response pathways in cell suspensions and plantlets of Arabidopsis thaliana: activation by hypoosmolarity and negative role in hyperosmolarity tolerance. FEBS Lett *574*, 42-48.

Duan, L., Dietrich, D., Ng, C.H., Chan, P.M., Bhalerao, R., Bennett, M.J., and Dinneny, J.R. (2013). Endodermal ABA signaling promotes lateral root quiescence during salt stress in Arabidopsis seedlings. Plant Cell *25*, 324-341.

Dubrovsky, J.G., Sauer, M., Napsucialy-Mendivil, S., Ivanchenko, M.G., Friml, J., Shishkova, S., Celenza, J., and Benkova, E. (2008). Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. Proc Natl Acad Sci U S A *105*, 8790-8794.

Dyer, N.P., Shahrezaei, V., and Hebenstreit, D. (2019). LiBiNorm: an htseq-count analogue with improved normalisation of Smart-seq2 data and library preparation diagnostics. PeerJ *7*.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A., and Liotta, L.A. (1996). Laser capture microdissection. Science *274*, 998-1001.

Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kalo, P., and Kiss, G.B. (2002). A receptor kinase gene regulating symbiotic nodule development. Nature *417*, 962-966.

Esseling, J.J., Lhuissier, F.G., and Emons, A.M. (2003). Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. Plant Physiol *132*, 1982-1988.

Eulgem, T., and Somssich, I.E. (2007). Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol *10*, 366-371.

Fan, S.C., Lin, C.S., Hsu, P.K., Lin, S.H., and Tsay, Y.F. (2009). The Arabidopsis nitrate transporter *NRT1.7,* expressed in phloem, is responsible for source-to-sink remobilization of nitrate. Plant Cell *21*, 2750-2761.

Ferguson, B.J., Li, D., Hastwell, A.H., Reid, D.E., Li, Y., Jackson, S.A., and Gresshoff, P.M. (2014). The soybean (*Glycine max*) nodulation-suppressive CLE peptide, *GmRIC1*, functions interspecifically in common white bean (*Phaseolus vulgaris*), but not in a supernodulating line mutated in the receptor *PvNARK*. Plant Biotechnol J *12*, 1085-1097.

Ferguson, B.J., Mens, C., Hastwell, A.H., Zhang, M., Su, H., Jones, C.H., Chu, X., and Gresshoff, P.M. (2019). Legume nodulation: The host controls the party. Plant Cell Environ *42*, 41-51.

Fisher, R.F., and Long, S.R. (1992). Rhizobium--plant signal exchange. Nature *357*, 655-660.

Fliegmann, J., and Bono, J.J. (2015). Lipo-chitooligosaccharidic nodulation factors and their perception by plant receptors. Glycoconj J *32*, 455-464.

Formey, D., Sallet, E., Lelandais-Briere, C., Ben, C., Bustos-Sanmamed, P., Niebel, A., Frugier, F., Combier, J.P., Debelle, F., Hartmann, C*., et al.* (2014). The small RNA diversity from *Medicago truncatula* roots under biotic interactions evidences the environmental plasticity of the miRNAome. Genome Biol *15*, 457.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., and Putterill, J. (1999). *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. EMBO J *18*, 4679-4688.

Fraysse, N., Couderc, F., and Poinsot, V. (2003). Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. Eur J Biochem *270*, 1365-1380.

Frerichs, A., Engelhorn, J., Altmuller, J., Gutierrez-Marcos, J., and Werr, W. (2019). Specific chromatin changes mark lateral organ founder cells in the *Arabidopsis thaliana* inflorescence meristem. J Exp Bot.

Fujikake, H., Yamazaki, A., Ohtake, N., Sueyoshi, K., Matsuhashi, S., Ito, T., Mizuniwa, C., Kume, T., Hashimoto, S., Ishioka, N.S.*, et al.* (2003). Quick and reversible inhibition of soybean root nodule growth by nitrate involves a decrease in sucrose supply to nodules. J Exp Bot *54*, 1379-1388.

Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M. (2002). Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT*/*IAA14* gene of Arabidopsis. Plant J *29*, 153-168.

Gage, D.J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev *68*, 280-300.

Galvan-Ampudia, C.S., and Testerink, C. (2011). Salt stress signals shape the plant root. Curr Opin Plant Biol *14*, 296-302.

Gamas, P., Brault, M., Jardinaud, M.F., and Frugier, F. (2017). Cytokinins in Symbiotic Nodulation: When, Where, What For? Trends Plant Sci *22*, 792-802.

Gan, Y., Filleur, S., Rahman, A., Gotensparre, S., and Forde, B.G. (2005). Nutritional regulation of *ANR1* and other root-expressed MADS-box genes in *Arabidopsis thaliana*. Planta *222*, 730-742.

Gao, M., Liu, J., Bi, D., Zhang, Z., Cheng, F., Chen, S., and Zhang, Y. (2008). *MEKK1*, *MKK1*/*MKK2* and *MPK4* function together in a mitogen-activated protein kinase cascade to regulate innate immunity in plants. Cell Res *18*, 1190-1198.

Gao, X., Chen, X., Lin, W., Chen, S., Lu, D., Niu, Y., Li, L., Cheng, C., McCormack, M., Sheen, J.*, et al.* (2013). Bifurcation of Arabidopsis *NLR* immune signaling via Ca^{2+}-dependent protein kinases. PLoS Pathog *9*, e1003127.

Geng, Y., Wu, R., Wee, C.W., Xie, F., Wei, X., Chan, P.M., Tham, C., Duan, L., and Dinneny, J.R. (2013). A spatio-temporal understanding of growth regulation during the salt stress response in Arabidopsis. Plant Cell *25*, 2132-2154.

Gibbs, D.J., and Coates, J.C. (2014). *AtMYB93* is an endodermis-specific transcriptional regulator of lateral root development in Arabidopsis. Plant Signal Behav *9*, e970406.

Gibbs, D.J., Voss, U., Harding, S.A., Fannon, J., Moody, L.A., Yamada, E., Swarup, K., Nibau, C., Bassel, G.W., Choudhary, A.*, et al.* (2014). *AtMYB93* is a novel negative regulator of lateral root development in Arabidopsis. New Phytol *203*, 1194-1207.

Gibson, K.E., Kobayashi, H., and Walker, G.C. (2008). Molecular determinants of a symbiotic chronic infection. Annu Rev Genet *42*, 413-441.

Gifford, M.L., Dean, A., Gutierrez, R.A., Coruzzi, G.M., and Birnbaum, K.D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci U S A *105*, 803-808.

Glawischnig, E. (2007). Camalexin. Phytochemistry *68*, 401-406.

Gleason, C., Chaudhuri, S., Yang, T., Munoz, A., Poovaiah, B.W., and Oldroyd, G.E. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. Nature *441*, 1149-1152.

Gomez-Gomez, L., and Boller, T. (2000). *FLS2*: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Mol Cell *5*, 1003-1011.

Gomez-Gomez, L., Felix, G., and Boller, T. (1999). A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. Plant J *18*, 277-284.

Gonzalez-Arzola, K., Diaz-Quintana, A., Rivero-Rodriguez, F., Velazquez-Campoy, A., De la Rosa, M.A., and Diaz-Moreno, I. (2017). Histone chaperone activity of *Arabidopsis thaliana NRP1* is blocked by cytochrome c. Nucleic Acids Res *45*, 2150-2165.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N.*, et al.* (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res *40*, D1178-1186.

Goyal, S., Lambert, C., Cluzet, S., Mérillon, J.M., and Ramawat, K.G. (2012). Secondary Metabolites and Plant Defence. In Plant Defence: Biological Control, pp. 109-138.

Gronlund, J.T., Eyres, A., Kumar, S., Buchanan-Wollaston, V., and Gifford, M.L. (2012). Cell specific analysis of Arabidopsis leaves using fluorescence activated cell sorting. J Vis Exp.

Gutierrez, L., Mongelard, G., Flokova, K., Pacurar, D.I., Novak, O., Staswick, P., Kowalczyk, M., Pacurar, M., Demailly, H., Geiss, G.*, et al.* (2012). Auxin controls Arabidopsis adventitious root initiation by regulating jasmonic acid homeostasis. Plant Cell *24*, 2515-2527.

Harrison, M.J., Dewbre, G.R., and Liu, J. (2002). A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. Plant Cell *14*, 2413-2429.

Hastwell, A.H., Corcilius, L., Williams, J.T., Gresshoff, P.M., Payne, R.J., and Ferguson, B.J. (2019). Triarabinosylation is required for nodulation-suppressive *CLE* peptides to systemically inhibit nodulation in *Pisum sativum*. Plant Cell Environ *42*, 188-197.

Hastwell, A.H., Gresshoff, P.M., and Ferguson, B.J. (2015). The structure and activity of nodulation-suppressing *CLE* peptide hormones of legumes. Functional Plant Biology *42*.

Hayashi, T., Shimoda, Y., Sato, S., Tabata, S., Imaizumi-Anraku, H., and Hayashi, M. (2014). Rhizobial infection does not require cortical expression of upstream common symbiosis genes responsible for the induction of Ca^{2+} spiking. Plant J *77*, 146-159.

Heard, N.A., Holmes, C.C., Stephens, D.A., Hand, D.J., and Dimopoulos, G. (2005). Bayesian coclustering of Anopheles gene expression time series: study of immune defense response to multiple experimental challenges. Proc Natl Acad Sci U S A *102*, 16939-16944.

Heidstra, R., Welch, D., and Scheres, B. (2004). Mosaic analyses using marked activation and deletion clones dissect Arabidopsis *SCARECROW* action in asymmetric cell division. Genes Dev *18*, 1964-1969.

Heiman, M., Kulicke, R., Fenster, R.J., Greengard, P., and Heintz, N. (2014). Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). Nat Protoc *9*, 1282-1291.

Helariutta, Y., Fukaki, H., Wysocka-Diller, J., Nakajima, K., Jung, J., Sena, G., Hauser, M.T., and Benfey, P.N. (2000). The *SHORT-ROOT* gene controls radial patterning of the Arabidopsis root through radial signaling. Cell *101*, 555-567.

Hematy, K., Cherk, C., and Somerville, S. (2009). Host-pathogen warfare at the plant cell wall. Curr Opin Plant Biol *12*, 406-413.

Herrbach, V., Rembliere, C., Gough, C., and Bensmihen, S. (2014). Lateral root formation and patterning in *Medicago truncatula*. J Plant Physiol *171*, 301-310.

Hirsch, S., Kim, J., Munoz, A., Heckmann, A.B., Downie, J.A., and Oldroyd, G.E. (2009). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. Plant Cell *21*, 545-557.

Horvath, B., Domonkos, A., Kereszt, A., Szucs, A., Abraham, E., Ayaydin, F., Boka, K., Chen, Y., Chen, R., Murray, J.D*., et al.* (2015). Loss of the nodule-specific cysteine rich peptide, *NCR169*, abolishes symbiotic nitrogen fixation in the *Medicago truncatula dnf7* mutant. Proc Natl Acad Sci U S A *112*, 15232-15237.

Hsu, P.Y., and Harmer, S.L. (2014). Wheels within wheels: the plant circadian system. Trends Plant Sci *19*, 240-249.

Huala, E., Dickerman, A.W., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M., Huang, W*., et al.* (2001). The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res *29*, 102-105.

Huang, N.C., Liu, K.H., Lo, H.J., and Tsay, Y.F. (1999). Cloning and functional characterization of an Arabidopsis nitrate transporter gene that encodes a constitutive component of low-affinity uptake. Plant Cell *11*, 1381-1392.

Huang, Q., and Allen, C. (1997). An exo-poly-alpha-D-galacturonosidase, *PehB*, is required for wild-type virulence of *Ralstonia solanacearum*. J Bacteriol *179*, 7369-7378.

Huffaker, A., Pearce, G., and Ryan, C.A. (2006). An endogenous peptide signal in Arabidopsis activates components of the innate immune response. Proc Natl Acad Sci U S A *103*, 10098-10103.

Hunt, S., and Layzell, D.B. (1993). Gas exchange of legume nodules and the regulation of nitrogenase actibvity. Annual Review of Plant Physiology and Plant Molecular Biology *44*, 483-511.

Ibanez, F., Wall, L., and Fabra, A. (2017). Starting points in plant-bacteria nitrogen-fixing symbioses: intercellular invasion of the roots. J Exp Bot *68*, 1905-1918.

Ichimura, K., Casais, C., Peck, S.C., Shinozaki, K., and Shirasu, K. (2006). MEKK1 is required for *MPK4* activation and regulates tissue-specific and temperature-dependent cell death in Arabidopsis. J Biol Chem *281*, 36969-36976.

Ivanova, K.A., Tsyganova, A.V., Brewin, N.J., Tikhonovich, I.A., and Tsyganov, V.E. (2015). Induction of host defences by Rhizobium during ineffective nodulation of pea (*Pisum sativum* L.) carrying symbiotically defective mutations *sym40* (*PsEFD*), *sym33* (*PsIPD3*/*PsCYCLOPS*) and *sym42*. Protoplasma.

Iyer-Pascuzzi, A.S., Jackson, T., Cui, H., Petricka, J.J., Busch, W., Tsukagoshi, H., and Benfey, P.N. (2011). Cell identity regulators link development and stress responses in the Arabidopsis root. Dev Cell *21*, 770-782.

Jacob, F., Vernaldi, S., and Maekawa, T. (2013). Evolution and Conservation of Plant NLR Functions. Front Immunol *4*, 297.

Javot, H., Penmetsa, R.V., Terzaghi, N., Cook, D.R., and Harrison, M.J. (2007). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci U S A *104*, 1720-1725.

Jones, K.M., Sharopova, N., Lohar, D.P., Zhang, J.Q., VandenBosch, K.A., and Walker, G.C. (2008). Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an exopolysaccharide-deficient mutant. Proc Natl Acad Sci U S A *105*, 704-709.

Journet, E.P., El-Gachtouli, N., Vernoud, V., de Billy, F., Pichon, M., Dedieu, A., Arnould, C., Morandi, D., Barker, D.G., and Gianinazzi-Pearson, V. (2001). *Medicago truncatula ENOD11*: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. Mol Plant Microbe Interact *14*, 737-748.

Kakar, K., Wandrey, M., Czechowski, T., Gaertner, T., Scheible, W.R., Stitt, M., Torres-Jerez, I., Xiao, Y., Redman, J.C., Wu, H.C*., et al.* (2008). A community

resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in *Medicago truncatula*. Plant Methods *4*, 18.

Kalo, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J.*, et al.* (2005). Nodulation signaling in legumes requires *NSP2*, a member of the GRAS family of transcriptional regulators. Science *308*, 1786-1789.

Kanamori, N., Madsen, L.H., Radutoiu, S., Frantescu, M., Quistgaard, E.M., Miwa, H., Downie, J.A., James, E.K., Felle, H.H., Haaning, L.L.*, et al.* (2006). A nucleoporin is required for induction of Ca^{2+} spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci U S A *103*, 359-364.

Kannan, P., Pandey, D., Gupta, A.K., Punetha, H., Taj, G., and Kumar, A. (2012). Expression analysis of *MAP2K9* and *MAPK6* during pathogenesis of Alternaria blight in *Arabidopsis thaliana* ecotype Columbia. Mol Biol Rep *39*, 4439-4444.

Kao, C.C., Barlow, E., and Sequeira, L. (1992). Extracellular polysaccharide is required for wild-type virulence of *Pseudomonas solanacearum*. J Bacteriol *174*, 1068-1071.

Kaschuk, G., Kuyper, T.W., Leffelaar, P.A., Hungria, M., and Giller, K.E. (2009). Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular mycorrhizal symbioses? Soil Biology and Biochemistry *41*, 1233-1244.

Kauffmann, A., Gentleman, R., and Huber, W. (2009). arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. Bioinformatics *25*, 415-416.

Kawaharada, Y., James, E.K., Kelly, S., Sandal, N., and Stougaard, J. (2017). The Ethylene Responsive Factor Required for Nodulation 1 (*ERN1*) Transcription Factor Is Required for Infection-Thread Formation in *Lotus japonicus*. Mol Plant Microbe Interact *30*, 194-204.

Kesarwani, M., Yoo, J., and Dong, X. (2007). Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in Arabidopsis. Plant Physiol *144*, 336-346.

Kiba, T., Feria-Bourrellier, A.B., Lafouge, F., Lezhneva, L., Boutet-Mercey, S., Orsel, M., Brehaut, V., Miller, A., Daniel-Vedele, F., Sakakibara, H.*, et al.* (2012). The Arabidopsis nitrate transporter *NRT2.4* plays a double role in roots and shoots of nitrogen-starved plants. Plant Cell *24*, 245-258.

Kiba, T., Inaba, J., Kudo, T., Ueda, N., Konishi, M., Mitsuda, N., Takiguchi, Y., Kondou, Y., Yoshizumi, T., Ohme-Takagi, M.*, et al.* (2018). Repression of Nitrogen Starvation Responses by Members of the Arabidopsis GARP-Type Transcription Factor *NIGT1*/*HRS1* Subfamily. Plant Cell *30*, 925-945.

Kiers, E.T., Rousseau, R.A., West, S.A., and Denison, R.F. (2003). Host sanctions and the legume-rhizobium mutualism. Nature *425*, 78-81.

Kim, H.S., Desveaux, D., Singer, A.U., Patel, P., Sondek, J., and Dangl, J.L. (2005a). The *Pseudomonas syringae* effector *AvrRpt2* cleaves its C-terminally acylated target, *RIN4*, from Arabidopsis membranes to block *RPM1* activation. Proc Natl Acad Sci U S A *102*, 6496-6501.

Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L., and Mackey, D. (2005b). Two *Pseudomonas syringae* type III effectors inhibit *RIN4*-regulated basal defense in Arabidopsis. Cell *121*, 749-759.

Kohorn, B.D., and Kohorn, S.L. (2012). The cell wall-associated kinases, WAKs, as pectin receptors. Front Plant Sci *3*, 88.

Konishi, M., and Yanagisawa, S. (2010). Identification of a nitrate-responsive cis-element in the Arabidopsis *NIR1* promoter defines the presence of multiple cis-regulatory elements for nitrogen response. Plant J *63*, 269-282.

Konishi, M., and Yanagisawa, S. (2011). The regulatory region controlling the nitrate-responsive expression of a nitrate reductase gene, *NIA1*, in Arabidopsis. Plant Cell Physiol *52*, 824-836.

Konishi, M., and Yanagisawa, S. (2013). Arabidopsis *NIN*-like transcription factors have a central role in nitrate signalling. Nat Commun *4*, 1617.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics *23*, 1289-1291.

Kotur, Z., Mackenzie, N., Ramesh, S., Tyerman, S.D., Kaiser, B.N., and Glass, A.D. (2012). Nitrate transport capacity of the *Arabidopsis thaliana* NRT2 family members and their interactions with *AtNAR2.1*. New Phytol *194*, 724-731.

Kouchi, H., Shimomura, K., Hata, S., Hirota, A., Wu, G.J., Kumagai, H., Tajima, S., Suganuma, N., Suzuki, A., Aoki, T.*, et al.* (2004). Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. DNA Res *11*, 263-274.

Krapp, A., David, L.C., Chardin, C., Girin, T., Marmagne, A., Leprince, A.S., Chaillou, S., Ferrario-Mery, S., Meyer, C., and Daniel-Vedele, F. (2014). Nitrate transport and signalling in Arabidopsis. J Exp Bot *65*, 789-798.

Krasileva, K.V., Dahlbeck, D., and Staskawicz, B.J. (2010). Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. Plant Cell *22*, 2444-2458.

Krol, E., Mentzel, T., Chinchilla, D., Boller, T., Felix, G., Kemmerling, B., Postel, S., Arents, M., Jeworutzki, E., Al-Rasheid, K.A.*, et al.* (2010). Perception of the

Arabidopsis danger signal peptide 1 involves the pattern recognition receptor *AtPEPR1* and its close homologue *AtPEPR2*. J Biol Chem *285*, 13471-13479.

Krouk, G., Crawford, N.M., Coruzzi, G.M., and Tsay, Y.F. (2010a). Nitrate signaling: adaptation to fluctuating environments. Curr Opin Plant Biol *13*, 266-273.

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., Hoyerova, K., Tillard, P., Leon, S., Ljung, K.*, et al.* (2010b). Nitrate-regulated auxin transport by *NRT1.1* defines a mechanism for nutrient sensing in plants. Dev Cell *18*, 927-937.

Krouk, G., Mirowski, P., LeCun, Y., Shasha, D.E., and Coruzzi, G.M. (2010c). Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. Genome Biol *11*, R123.

Krusell, L., Madsen, L.H., Sato, S., Aubert, G., Genua, A., Szczyglowski, K., Duc, G., Kaneko, T., Tabata, S., de Bruijn, F.*, et al.* (2002). Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature *420*, 422-426.

Lagunas, B., Achom, M., Bonyadi-Pour, R., Pardal, A.J., Richmond, B.L., Sergaki, C., Vazquez, S., Schafer, P., Ott, S., Hammond, J.*, et al.* (2019). Regulation of Resource Partitioning Coordinates Nitrogen and Rhizobia Responses and Autoregulation of Nodulation in *Medicago truncatula*. Mol Plant *12*, 833-846.

Lagunas, B., Schafer, P., and Gifford, M.L. (2015). Housing helpful invaders: the evolutionary and molecular architecture underlying plant root-mutualist microbe interactions. J Exp Bot *66*, 2177-2186.

Laplaze, L., Benkova, E., Casimiro, I., Maes, L., Vanneste, S., Swarup, R., Weijers, D., Calvo, V., Parizot, B., Herrera-Rodriguez, M.B.*, et al.* (2007). Cytokinins act directly on lateral root founder cells to inhibit root initiation. Plant Cell *19*, 3889-3900.

Laskowski, M.J., Williams, M.E., Nusbaum, H.C., and Sussex, I.M. (1995). Formation of lateral root meristems is a two-stage process. Development *121*, 3303-3310.

Li, W., Wang, Y., Okamoto, M., Crawford, N.M., Siddiqi, M.Y., and Glass, A.D. (2007). Dissection of the *AtNRT2.1*:*AtNRT2.2* inducible high-affinity nitrate transporter gene cluster. Plant Physiol *143*, 425-433.

Liang, J., Klingl, A., Lin, Y.Y., Boul, E., Thomas-Oates, J., and Marin, M. (2019). A subcompatible rhizobium strain reveals infection duality in Lotus. J Exp Bot *70*, 1903-1913.

Liang, Y., Cao, Y., Tanaka, K., Thibivilliers, S., Wan, J., Choi, J., Kang, C., Qiu, J., and Stacey, G. (2013). Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. Science *341*, 1384-1387.

Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R.J., Bilgin, D.D., Radwan, O., Neece, D.J., Clough, S.J., May, G.D.*, et al.* (2010). Complete

transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. Plant Physiol *152*, 541-552.

Lin, M.H., Gresshoff, P.M., and Ferguson, B.J. (2012). Systemic regulation of soybean nodulation by acidic growth conditions. Plant Physiol *160*, 2028-2039.

Lin, S.H., Kuo, H.F., Canivenc, G., Lin, C.S., Lepetit, M., Hsu, P.K., Tillard, P., Lin, H.L., Wang, Y.Y., Tsai, C.B*., et al.* (2008). Mutation of the Arabidopsis *NRT1.5* nitrate transporter causes defective root-to-shoot nitrate transport. Plant Cell *20*, 2514-2528.

Linkohr, B.I., Williamson, L.C., Fitter, A.H., and Leyser, H.M. (2002). Nitrate and phosphate availability and distribution have different effects on root system architecture of Arabidopsis. Plant J *29*, 751-760.

Little, D.Y., Rao, H., Oliva, S., Daniel-Vedele, F., Krapp, A., and Malamy, J.E. (2005). The putative high-affinity nitrate transporter *NRT2.1* represses lateral root initiation in response to nutritional cues. Proc Natl Acad Sci U S A *102*, 13693-13698.

Liu, C.W., and Murray, J.D. (2016). The Role of Flavonoids in Nodulation Host-Range Specificity: An Update. Plants (Basel) *5*.

Liu, J., Elmore, J.M., and Coaker, G. (2009a). Investigating the functions of the *RIN4* protein complex during plant innate immune responses. Plant Signal Behav *4*, 1107-1110.

Liu, J., Elmore, J.M., Fuglsang, A.T., Palmgren, M.G., Staskawicz, B.J., and Coaker, G. (2009b). *RIN4* functions with plasma membrane H^+-ATPases to regulate stomatal apertures during pathogen attack. PLoS Biol *7*, e1000139.

Liu, K.H., and Tsay, Y.F. (2003). Switching between the two action modes of the dual-affinity nitrate transporter *CHL1* by phosphorylation. EMBO J *22*, 1005-1013.

Liu, Y., Hassan, S., Kidd, B.N., Garg, G., Mathesius, U., Singh, K.B., and Anderson, J.P. (2017). Ethylene Signaling Is Important for Isoflavonoid-Mediated Resistance to *Rhizoctonia solani* in Roots of *Medicago truncatula*. Mol Plant Microbe Interact *30*, 691-700.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lohar, D.P., Sharopova, N., Endre, G., Penuela, S., Samac, D., Town, C., Silverstein, K.A., and VandenBosch, K.A. (2006). Transcript analysis of early nodulation events in *Medicago truncatula*. Plant Physiol *140*, 221-234.

Long, S.R. (1996). Rhizobium symbiosis: nod factors in perspective. Plant Cell *8*, 1885-1898.

Lopez-Gomez, M., Sandal, N., Stougaard, J., and Boller, T. (2012). Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. J Exp Bot *63*, 393-401.

Loque, D., Yuan, L., Kojima, S., Gojon, A., Wirth, J., Gazzarrini, S., Ishiyama, K., Takahashi, H., and von Wiren, N. (2006). Additive contribution of *AMT1;1* and *AMT1;3* to high-affinity ammonium uptake across the plasma membrane of nitrogen-deficient Arabidopsis roots. Plant J *48*, 522-534.

Lorang, J.M., Sweat, T.A., and Wolpert, T.J. (2007). Plant disease susceptibility conferred by a "resistance" gene. Proc Natl Acad Sci U S A *104*, 14861-14866.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*.

Lowe-Power, T.M., Khokhani, D., and Allen, C. (2018). How *Ralstonia solanacearum* Exploits and Thrives in the Flowing Plant Xylem Environment. Trends Microbiol *26*, 929-942.

Lucas, M., Swarup, R., Paponov, I.A., Swarup, K., Casimiro, I., Lake, D., Peret, B., Zappala, S., Mairhofer, S., Whitworth, M.*, et al.* (2011). Short-Root regulates primary, lateral, and adventitious root development in Arabidopsis. Plant Physiol *155*, 384-398.

Lynch, J. (1995). Root Architecture and Plant Productivity. Plant Physiol *109*, 7-13.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R., and Dangl, J.L. (2003). Arabidopsis *RIN4* is a target of the type III virulence effector *AvrRpt2* and modulates *RPS2*-mediated resistance. Cell *112*, 379-389.

Mackey, D., Holt, B.F., 3rd, Wiig, A., and Dangl, J.L. (2002). *RIN4* interacts with *Pseudomonas syringae* type III effector molecules and is required for *RPM1*-mediated resistance in Arabidopsis. Cell *108*, 743-754.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M.*, et al.* (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202-1214.

Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N.*, et al.* (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature *425*, 637-640.

Magori, S., and Kawaguchi, M. (2009). Long-distance control of nodulation: molecules and models. Mol Cells *27*, 129-134.

Magori, S., Oka-Kira, E., Shibata, S., Umehara, Y., Kouchi, H., Hase, Y., Tanaka, A., Sato, S., Tabata, S., and Kawaguchi, M. (2009). *Too much love*, a root regulator

associated with the long-distance control of nodulation in *Lotus japonicus*. Mol Plant Microbe Interact *22*, 259-268.

Mah, K.M., Uppalapati, S.R., Tang, Y., Allen, S., and Shuai, B. (2012). Gene expression profiling of *Macrophomina phaseolina* infected *Medicago truncatula* roots reveals a role for auxin in plant tolerance against the charcoal rot pathogen. Physiological and Molecular Plant Pathology *79*, 21-30.

Malamy, J.E., and Ryan, K.S. (2001). Environmental regulation of lateral root initiation in Arabidopsis. Plant Physiol *127*, 899-909.

Manchanda, G., and Garg, N. (2008). Salinity and its effects on the functional biology of legumes. Acta Physiologiae Plantarum *30*, 595-618.

Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., and Zhang, S. (2011). Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. Plant Cell *23*, 1639-1653.

Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., Timmers, T., Poinsot, V., Gilbert, L.B., Heeb, P*., et al.* (2010). Experimental evolution of a plant pathogen into a legume symbiont. PLoS Biol *8*, e1000280.

Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., Cruveiller, S., Jauneau, A., Capela, D., and Masson-Boivin, C. (2017). Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. Mol Ecol *26*, 1818-1831.

Marchetti, M., Jauneau, A., Capela, D., Remigi, P., Gris, C., Batut, J., and Masson-Boivin, C. (2014). Shaping bacterial symbiosis with legumes by experimental evolution. Mol Plant Microbe Interact *27*, 956-964.

Marchive, C., Roudier, F., Castaings, L., Brehaut, V., Blondet, E., Colot, V., Meyer, C., and Krapp, A. (2013). Nuclear retention of the transcription factor *NLP7* orchestrates the early response to nitrate in plants. Nat Commun *4*, 1713.

Marsh, J.F., Rakocevic, A., Mitra, R.M., Brocard, L., Sun, J., Eschstruth, A., Long, S.R., Schultze, M., Ratet, P., and Oldroyd, G.E. (2007). *Medicago truncatula NIN* is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. Plant Physiol *144*, 324-335.

Mauch, F., Mauch-Mani, B., and Boller, T. (1988). Antifungal Hydrolases in Pea Tissue : II. Inhibition of Fungal Growth by Combinations of Chitinase and beta-1,3-Glucanase. Plant Physiol *88*, 936-942.

Mbengue, M., Camut, S., de Carvalho-Niebel, F., Deslandes, L., Froidure, S., Klaus-Heisen, D., Moreau, S., Rivas, S., Timmers, T., Herve, C*., et al.* (2010). The *Medicago truncatula* E3 ubiquitin ligase *PUB1* interacts with the *LYK3* symbiotic receptor and negatively regulates infection and nodulation. Plant Cell *22*, 3474-3488.

McHale, L., Tan, X., Koehl, P., and Michelmore, R.W. (2006). Plant NBS-LRR proteins: adaptable guards. Genome Biol *7*, 212.

Medici, A., and Krouk, G. (2014). The primary nitrate response: a multifaceted signalling pathway. J Exp Bot *65*, 5567-5576.

Medici, A., Marshall-Colon, A., Ronzier, E., Szponarski, W., Wang, R., Gojon, A., Crawford, N.M., Ruffel, S., Coruzzi, G.M., and Krouk, G. (2015). *AtNIGT1/HRS1* integrates nitrate and phosphate signals at the Arabidopsis root tip. Nat Commun *6*, 6274.

Messinese, E., Mun, J.H., Yeun, L.H., Jayaraman, D., Rouge, P., Barre, A., Lougnon, G., Schornack, S., Bono, J.J., Cook, D.R*., et al.* (2007). A novel nuclear protein interacts with the symbiotic *DMI3* calcium- and calmodulin-dependent protein kinase of *Medicago truncatula*. Mol Plant Microbe Interact *20*, 912-921.

Middleton, P.H., Jakab, J., Penmetsa, R.V., Starker, C.G., Doll, J., Kalo, P., Prabhu, R., Marsh, J.F., Mitra, R.M., Kereszt, A*., et al.* (2007). An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. Plant Cell *19*, 1221-1234.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H., and Shibuya, N. (2007). *CERK1*, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. Proc Natl Acad Sci U S A *104*, 19613-19618.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H.R., Carre, I.A., and Coupland, G. (2002). *LHY* and *CCA1* are partially redundant genes required to maintain circadian rhythms in Arabidopsis. Dev Cell *2*, 629-641.

Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M.A., Limpens, E., Gadella, T.W., and Bisseling, T. (2014). Nod factor receptors form heteromeric complexes and are essential for intracellular infection in Medicago nodules. Plant Cell *26*, 4188-4199.

Moreau, S., Fromentin, J., Vailleau, F., Vernie, T., Huguet, S., Balzergue, S., Frugier, F., Gamas, P., and Jardinaud, M.F. (2014). The symbiotic transcription factor *MtEFD* and cytokinins are positively acting in the *Medicago truncatula* and *Ralstonia solanacearum* pathogenic interaction. New Phytol *201*, 1343-1357.

Mortier, V., Den Herder, G., Whitford, R., Van de Velde, W., Rombauts, S., D'Haeseleer, K., Holsters, M., and Goormachtig, S. (2010). *CLE* peptides control *Medicago truncatula* nodulation locally and systemically. Plant Physiol *153*, 222-237.

Mortier, V., Holsters, M., and Goormachtig, S. (2012). Never too many? How legumes control nodule numbers. Plant Cell Environ *35*, 245-258.

Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T.W., Belcher, H., Yativ, M., Brady, S.M., Benfey, P.N., and Aharoni, A. (2013). High-resolution metabolic mapping of cell types in plant roots. Proc Natl Acad Sci U S A *110*, E1232-1241.

Mukhtar, M.S., Deslandes, L., Auriac, M.C., Marco, Y., and Somssich, I.E. (2008). The Arabidopsis transcription factor *WRKY27* influences wilt disease symptom development caused by *Ralstonia solanacearum*. Plant J *56*, 935-947.

Munoz, V., Ibanez, F., Tordable, M., Megias, M., and Fabra, A. (2015). Role of reactive oxygen species generation and Nod factors during the early symbiotic interaction between bradyrhizobia and peanut, a legume infected by crack entry. J Appl Microbiol *118*, 182-192.

Nagata, M., Murakami, E., Shimoda, Y., Shimoda-Sasakura, F., Kucho, K., Suzuki, A., Abe, M., Higashi, S., and Uchiumi, T. (2008). Expression of a class 1 hemoglobin gene and production of nitric oxide in response to symbiotic and pathogenic bacteria in *Lotus japonicus*. Mol Plant Microbe Interact *21*, 1175-1183.

Nakagawa, T., and Kawaguchi, M. (2006). Shoot-applied MeJA Suppresses Root Nodulation in *Lotus japonicus*. Plant and Cell Physiology *47*, 176-180.

Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T., and Jones, J.D. (2004). The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. Plant Physiol *135*, 1113-1128.

Nawy, T., Lee, J.Y., Colinas, J., Wang, J.Y., Thongrod, S.C., Malamy, J.E., Birnbaum, K., and Benfey, P.N. (2005). Transcriptional profile of the Arabidopsis root quiescent center. Plant Cell *17*, 1908-1925.

Nishida, H., Handa, Y., Tanaka, S., Suzaki, T., and Kawaguchi, M. (2016). Expression of the *CLE-RS3* gene suppresses root nodulation in *Lotus japonicus*. J Plant Res *129*, 909-919.

Nishimura, R., Hayashi, M., Wu, G.J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S., Akao, S., Ohmori, M., Nagasawa, M*., et al.* (2002). *HAR1* mediates systemic regulation of symbiotic organ development. Nature *420*, 426-429.

O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. Cell *165*, 1280-1292.

Obertello, M., Krouk, G., Katari, M.S., Runko, S.J., and Coruzzi, G.M. (2010). Modeling the global effect of the basic-leucine zipper transcription factor 1 (*bZIP1*) on nitrogen and light regulation in Arabidopsis. BMC Syst Biol *4*, 111.

Okamoto, M., Kumar, A., Li, W., Wang, Y., Siddiqi, M.Y., Crawford, N.M., and Glass, A.D. (2006). High-affinity nitrate transport in roots of Arabidopsis depends on expression of the *NAR2*-like gene *AtNRT3.1*. Plant Physiol *140*, 1036-1046.

Okamoto, S., and Kawaguchi, M. (2015). Shoot *HAR1* mediates nitrate inhibition of nodulation in *Lotus japonicus*. Plant Signal Behav *10*, e1000138.

Okamoto, S., Ohnishi, E., Sato, S., Takahashi, H., Nakazono, M., Tabata, S., and Kawaguchi, M. (2009). Nod factor/nitrate-induced *CLE* genes that drive *HAR1*-mediated systemic regulation of nodulation. Plant Cell Physiol *50*, 67-77.

Okamoto, S., Shinohara, H., Mori, T., Matsubayashi, Y., and Kawaguchi, M. (2013). Root-derived *CLE* glycopeptides control nodulation by direct binding to *HAR1* receptor kinase. Nat Commun *4*, 2191.

Okazaki, S., Kaneko, T., Sato, S., and Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. Proc Natl Acad Sci U S A *110*, 17131-17136.

Oke, V., and Long, S.R. (1999). Bacteroid formation in the Rhizobium-legume symbiosis. Curr Opin Microbiol *2*, 641-646.

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tasaka, M. (2007). *ARF7* and *ARF19* regulate lateral root formation via direct activation of *LBD*/*ASL* genes in Arabidopsis. Plant Cell *19*, 118-130.

Oldroyd, G.E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol *11*, 252-263.

Omrane, S., and Chiurazzi, M. (2009). A variety of regulatory mechanisms are involved in the nitrogen-dependent modulation of the nodule organogenesis program in legume roots. Plant Signal Behav *4*, 1066-1068.

Omrane, S., Ferrarini, A., D'Apuzzo, E., Rogato, A., Delledonne, M., and Chiurazzi, M. (2009). Symbiotic competence in *Lotus japonicus* is affected by plant nitrogen status: transcriptomic identification of genes affected by a new signalling pathway. New Phytol *183*, 380-394.

Oono, R., Anderson, C.G., and Denison, R.F. (2011). Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. Proc Biol Sci *278*, 2698-2703.

Orman-Ligeza, B., Morris, E.C., Parizot, B., Lavigne, T., Babe, A., Ligeza, A., Klein, S., Sturrock, C., Xuan, W., Novak, O*., et al.* (2018). The Xerobranching Response Represses Lateral Root Formation When Roots Are Not in Contact with Water. Curr Biol *28*, 3165-3173 e3165.

Orosa-Puente, B., Leftley, N., von Wangenheim, D., Banda, J., Srivastava, A.K., Hill, K., Truskina, J., Bhosale, R., Morris, E., Srivastava, M*., et al.* (2018). Root branching toward water involves posttranslational modification of transcription factor *ARF7*. Science *362*, 1407-1410.

Orsel, M., Krapp, A., and Daniel-Vedele, F. (2002). Analysis of the NRT2 nitrate transporter family in Arabidopsis. Structure and gene expression. Plant Physiol *129*, 886-896.

Pan, H., and Wang, D. (2017). Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. Nat Plants *3*, 17048.

Panikulangara, T.J., Eggers-Schumacher, G., Wunderlich, M., Stransky, H., and Schoffl, F. (2004). *Galactinol synthase1*. A novel heat shock factor target gene responsible for heat-induced synthesis of raffinose family oligosaccharides in Arabidopsis. Plant Physiol *136*, 3148-3158.

Para, A., Li, Y., Marshall-Colon, A., Varala, K., Francoeur, N.J., Moran, T.M., Edwards, M.B., Hackley, C., Bargmann, B.O., Birnbaum, K.D*., et al.* (2014). Hit-and-run transcriptional control by *bZIP1* mediates rapid nutrient signaling in Arabidopsis. Proc Natl Acad Sci U S A *111*, 10371-10376.

Patriarca, E.J., Tate, R., Ferraioli, S., and Iaccarino, M. (2004). Organogenesis of legume root nodules. Int Rev Cytol *234*, 201-262.

Peck, M.C., Fisher, R.F., and Long, S.R. (2006). Diverse flavonoids stimulate *NodD1* binding to nod gene promoters in Sinorhizobium meliloti. J Bacteriol *188*, 5417-5427.

Peiter, E., Sun, J., Heckmann, A.B., Venkateshwaran, M., Riely, B.K., Otegui, M.S., Edwards, A., Freshour, G., Hahn, M.G., Cook, D.R*., et al.* (2007). The *Medicago truncatula DMI1* protein modulates cytosolic calcium signaling. Plant Physiol *145*, 192-203.

Peret, B., Larrieu, A., and Bennett, M.J. (2009). Lateral root emergence: a difficult birth. J Exp Bot *60*, 3637-3643.

Petersson, S.V., Johansson, A.I., Kowalczyk, M., Makoveychuk, A., Wang, J.Y., Moritz, T., Grebe, M., Benfey, P.N., Sandberg, G., and Ljung, K. (2009). An auxin gradient and maximum in the Arabidopsis root apex shown by high-resolution cell-specific analysis of IAA distribution and synthesis. Plant Cell *21*, 1659-1668.

Petricka, J.J., Schauer, M.A., Megraw, M., Breakfield, N.W., Thompson, J.W., Georgiev, S., Soderblom, E.J., Ohler, U., Moseley, M.A., Grossniklaus, U*., et al.* (2012). The protein expression landscape of the Arabidopsis root. Proc Natl Acad Sci U S A *109*, 6811-6818.

Petutschnig, E.K., Jones, A.M., Serazetdinova, L., Lipka, U., and Lipka, V. (2010). The lysin motif receptor-like kinase (LysM-RLK) *CERK1* is a major chitin-binding protein in *Arabidopsis thaliana* and subject to chitin-induced phosphorylation. J Biol Chem *285*, 28902-28911.

Pfund, C., Tans-Kersten, J., Dunning, F.M., Alonso, J.M., Ecker, J.R., Allen, C., and Bent, A.F. (2004). Flagellin is not a major defense elicitor in *Ralstonia solanacearum*

cells or extracts applied to *Arabidopsis thaliana*. Mol Plant Microbe Interact *17*, 696-706.

Pii, Y., Astegno, A., Peroni, E., Zaccardelli, M., Pandolfini, T., and Crimi, M. (2009). The *Medicago truncatula N5* gene encoding a root-specific lipid transfer protein is required for the symbiotic interaction with *Sinorhizobium meliloti*. Mol Plant Microbe Interact *22*, 1577-1587.

Pii, Y., Molesini, B., Masiero, S., and Pandolfini, T. (2012). The non-specific lipid transfer protein *N5* of *Medicago truncatula* is implicated in epidermal stages of rhizobium-host interaction. BMC Plant Biol *12*, 233.

Pitzschke, A., Schikora, A., and Hirt, H. (2009). MAPK cascade signalling networks in plant defence. Curr Opin Plant Biol *12*, 421-426.

Poueymiro, M., and Genin, S. (2009). Secreted proteins from *Ralstonia solanacearum*: a hundred tricks to kill a plant. Curr Opin Microbiol *12*, 44-52.

Price, P.A., Tanner, H.R., Dillon, B.A., Shabab, M., Walker, G.C., and Griffitts, J.S. (2015). Rhizobial peptidase *HrrP* cleaves host-encoded signaling peptides and mediates symbiotic compatibility. Proc Natl Acad Sci U S A *112*, 15244-15249.

Qiu, J.L., Fiil, B.K., Petersen, K., Nielsen, H.B., Botanga, C.J., Thorgrimsen, S., Palma, K., Suarez-Rodriguez, M.C., Sandbech-Clausen, S., Lichota, J.*, et al.* (2008). Arabidopsis *MAP kinase 4* regulates gene expression through transcription factor release in the nucleus. EMBO J *27*, 2214-2221.

Qutob, D., Kemmerling, B., Brunner, F., Kufner, I., Engelhardt, S., Gust, A.A., Luberacki, B., Seitz, H.U., Stahl, D., Rauhut, T.*, et al.* (2006). Phytotoxicity and innate immune responses induced by *Nep1*-like proteins. Plant Cell *18*, 3721-3744.

Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N.*, et al.* (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature *425*, 585-592.

Radutoiu, S., Madsen, L.H., Madsen, E.B., Jurkiewicz, A., Fukai, E., Quistgaard, E.M., Albrektsen, A.S., James, E.K., Thirup, S., and Stougaard, J. (2007). LysM domains mediate lipochitin-oligosaccharide recognition and Nfr genes extend the symbiotic host range. EMBO J *26*, 3923-3935.

Rafiqi, M., Bernoux, M., Ellis, J.G., and Dodds, P.N. (2009). In the trenches of plant pathogen recognition: Role of NB-LRR proteins. Semin Cell Dev Biol *20*, 1017-1024.

Rajeevkumar, S., Anunanthini, P., and Sathishkumar, R. (2015). Epigenetic silencing in transgenic plants. Front Plant Sci *6*, 693.

Rasmussen, M.W., Roux, M., Petersen, M., and Mundy, J. (2012). MAP Kinase Cascades in Arabidopsis Innate Immunity. Front Plant Sci *3*, 169.

Reid, D.E., Ferguson, B.J., Hayashi, S., Lin, Y.H., and Gresshoff, P.M. (2011). Molecular mechanisms controlling legume autoregulation of nodulation. Ann Bot *108*, 789-795.

Remans, T., Nacry, P., Pervent, M., Filleur, S., Diatloff, E., Mounier, E., Tillard, P., Forde, B.G., and Gojon, A. (2006). The Arabidopsis *NRT1.1* transporter participates in the signaling pathway triggering root colonization of nitrate-rich patches. Proc Natl Acad Sci U S A *103*, 19206-19211.

Ren, D., Liu, Y., Yang, K.Y., Han, L., Mao, G., Glazebrook, J., and Zhang, S. (2008). A fungal-responsive MAPK cascade regulates phytoalexin biosynthesis in Arabidopsis. Proc Natl Acad Sci U S A *105*, 5638-5643.

Rey, T., Bonhomme, M., Chatterjee, A., Gavrin, A., Toulotte, J., Yang, W., Andre, O., Jacquet, C., and Schornack, S. (2017). The Medicago truncatula GRAS protein *RAD1* supports arbuscular mycorrhiza symbiosis and *Phytophthora palmivora* susceptibility. J Exp Bot *68*, 5871-5881.

Rich, C., Reitz, M.U., Eichmann, R., Hermann, S., Jenkins, D.J., Kogel, K.-H., Esteban, E., Ott, S., and Schäfer, P. (2019). Cell type identity determines transcriptomic immune responses in *Arabidopsis thaliana*
roots.

Roden, L.C., and Ingle, R.A. (2009). Lights, rhythms, infection: the role of light and the circadian clock in determining the outcome of plant-pathogen interactions. Plant Cell *21*, 2546-2552.

Rogg, L.E., Lasswell, J., and Bartel, B. (2001). A gain-of-function mutation in *IAA28* suppresses lateral root development. Plant Cell *13*, 465-480.

Ross, A., Yamada, K., Hiruma, K., Yamashita-Yamada, M., Lu, X., Takano, Y., Tsuda, K., and Saijo, Y. (2014). The Arabidopsis PEPR pathway couples local and systemic plant immunity. EMBO J *33*, 62-75.

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J.C., Cattolico, L*., et al.* (2002). Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature *415*, 497-502.

Samuel, M.A., Mudgil, Y., Salt, J.N., Delmas, F., Ramachandran, S., Chilelli, A., and Goring, D.R. (2008). Interactions between the S-domain receptor kinases and *At*PUB-ARM E3 ubiquitin ligases suggest a conserved signaling pathway in Arabidopsis. Plant Physiol *147*, 2084-2095.

Santos, R., Herouart, D., Sigaud, S., Touati, D., and Puppo, A. (2001). Oxidative burst in Alfalfa-*Sinorhizobium meliloti* symbiotic interaction. Mol Plant Microbe Interact *14*, 86-89.

Scheidle, H., Gross, A., and Niehaus, K. (2005). The Lipid A substructure of the *Sinorhizobium meliloti* lipopolysaccharides is sufficient to suppress the oxidative burst in host plants. New Phytol *165*, 559-565.

Schell, M.A. (1987). Purification and Characterization of an Endoglucanase from *Pseudomonas solanacearum*. Appl Environ Microbiol *53*, 2237-2241.

Schlaman, H.R., Okker, R.J., and Lugtenberg, B.J. (1992). Regulation of nodulation gene expression by *NodD* in rhizobia. J Bacteriol *174*, 5177-5182.

Schnabel, E., Journet, E.P., de Carvalho-Niebel, F., Duc, G., and Frugoli, J. (2005). The Medicago truncatula *SUNN* gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. Plant Mol Biol *58*, 809-822.

Searle, I.R., Men, A.E., Laniya, T.S., Buzas, D.M., Iturbe-Ormaetxe, I., Carroll, B.J., and Gresshoff, P.M. (2003). Long-distance signaling in nodulation directed by a *CLAVATA1*-like receptor kinase. Science *299*, 109-112.

Singh, V.K., Munro, K., and Jia, Z. (2012). A novel calmodulin-beta-PIX interaction and its implication in receptor tyrosine kinase regulation. Cell Signal *24*, 1790-1796.

Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T. (2007). Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. Plant Physiol *145*, 183-191.

Smit, P., Raedts, J., Portyanko, V., Debelle, F., Gough, C., Bisseling, T., and Geurts, R. (2005). *NSP1* of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. Science *308*, 1789-1791.

Soyano, T., Hirakawa, H., Sato, S., Hayashi, M., and Kawaguchi, M. (2014). Nodule Inception creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. Proc Natl Acad Sci U S A *111*, 14607-14612.

Sprent, J.I. (2008). 60Ma of legume nodulation. What's new? What's changing? J Exp Bot *59*, 1081-1084.

Staal, J., and Dixelius, C. (2007). Tracing the ancient origins of plant innate immunity. Trends Plant Sci *12*, 334-342.

Stacey, G., McAlvin, C.B., Kim, S.-Y., Olivares, J., and José Soto, M. (2006). Effects of Endogenous Salicylic Acid on Nodulation in the Model Legumes *Lotus japonicus* and *Medicago truncatula*. Plant Physiology *141*, 1471-1481.

Stegle, O., Denby, K.J., Cooke, E.J., Wild, D.L., Ghahramani, Z., and Borgwardt, K.M. (2010). A robust Bayesian two-sample test for detecting intervals of differential gene expression in microarray time series. J Comput Biol *17*, 355-367.

Stougaard, J. (2000). Regulators and regulation of legume root nodule development. Plant Physiol *124*, 531-540.

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K.*, et al.* (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature *417*, 959-962.

Streeter, J.G. (2003). Effects of drought on nitrogen fixation in soybean root nodules. Plant Cell and Environment *26*, 1199-1204.

Subramanian, S., Stacey, G., and Yu, O. (2007). Distinct, crucial roles of flavonoids during legume nodulation. Trends Plant Sci *12*, 282-285.

Sun, J., Cardoza, V., Mitchell, D.M., Bright, L., Oldroyd, G., and Harris, J.M. (2006). Crosstalk between jasmonic acid, ethylene and Nod factor signalling allows integration of diverse inputs for regulation of nodulation. The Plant Journal *46*, 961-970.

Takahara, M., Magori, S., Soyano, T., Okamoto, S., Yoshida, C., Yano, K., Sato, S., Tabata, S., Yamaguchi, K., Shigenobu, S.*, et al.* (2013). Too much love, a novel Kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume-Rhizobium symbiosis. Plant Cell Physiol *54*, 433-447.

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K.L., Yandell, M., Gundlach, H.*, et al.* (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. BMC Genomics *15*, 312.

Tans-Kersten, J., Guan, Y., and Allen, C. (1998). *Ralstonia solanacearum* pectin methylesterase is required for growth on methylated pectin but not for bacterial wilt virulence. Appl Environ Microbiol *64*, 4918-4923.

Tans-Kersten, J., Huang, H., and Allen, C. (2001). *Ralstonia solanacearum* needs motility for invasive virulence on tomato. J Bacteriol *183*, 3597-3605.

Tatematsu, K., Kumagai, S., Muto, H., Sato, A., Watahiki, M.K., Harper, R.M., Liscum, E., and Yamamoto, K.T. (2004). *MASSUGU2* encodes Aux/*IAA19*, an auxin-regulated protein that functions together with the transcriptional activator *NPH4*/*ARF7* to regulate differential growth responses of hypocotyl and formation of lateral roots in *Arabidopsis thaliana*. Plant Cell *16*, 379-393.

Tellstrom, V., Usadel, B., Thimm, O., Stitt, M., Kuster, H., and Niehaus, K. (2007). The lipopolysaccharide of *Sinorhizobium meliloti* suppresses defense-associated gene expression in cell cultures of the host plant *Medicago truncatula*. Plant Physiol *143*, 825-837.

Terpolilli, J.J., O'Hara, G.W., Tiwari, R.P., Dilworth, M.J., and Howieson, J.G. (2008). The model legume *Medicago truncatula* A17 is poorly matched for N2 fixation with the sequenced microsymbiont *Sinorhizobium meliloti* 1021. New Phytol *179*, 62-66.

Thatcher, L.F., Williams, A.H., Garg, G., Buck, S.G., and Singh, K.B. (2016). Transcriptome analysis of the fungal pathogen *Fusarium oxysporum* f. sp.

medicaginis during colonisation of resistant and susceptible *Medicago truncatula* hosts identifies differential pathogenicity profiles and novel candidate effectors. BMC Genomics *17*, 860.

Tian, Q., and Reed, J.W. (1999). Control of auxin-regulated root development by the Arabidopsis thaliana *SHY2*/*IAA3* gene. Development *126*, 711-721.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Research *45*, W122-W129.

Timmers, A.C., Auriac, M.C., de Billy, F., and Truchet, G. (1998). Nod factor internalization and microtubular cytoskeleton changes occur concomitantly during nodule differentiation in Alfalfa. Development *125*, 339-349.

Timmers, A.C., Auriac, M.C., and Truchet, G. (1999). Refined analysis of early symbiotic steps of the Rhizobium-Medicago interaction in relationship with microtubular cytoskeleton rearrangements. Development *126*, 3617-3628.

Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L.H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A.S., Kawaguchi, M*., et al.* (2006). Deregulation of a Ca^{2+}/calmodulin-dependent kinase leads to spontaneous nodule development. Nature *441*, 1153-1156.

Turner, M., Jauneau, A., Genin, S., Tavella, M.J., Vailleau, F., Gentzbittel, L., and Jardinaud, M.F. (2009). Dissection of bacterial Wilt on *Medicago truncatula* revealed two type III secretion system effectors acting on root infection process and disease development. Plant Physiol *150*, 1713-1722.

Tworak, A., Urbanowicz, A., Podkowinski, J., Kurzynska-Kokorniak, A., Koralewska, N., and Figlerowicz, M. (2016). Six *Medicago truncatula* Dicer-like protein genes are expressed in plant cells and upregulated in nodules. Plant Cell Rep *35*, 1043-1052.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res *40*, e115.

Vailleau, F., Sartorel, E., Jardinaud, M.F., Chardon, F., Genin, S., Huguet, T., Gentzbittel, L., and Petitprez, M. (2007). Characterization of the interaction between the bacterial wilt pathogen *Ralstonia solanacearum* and the model legume plant *Medicago truncatula*. Mol Plant Microbe Interact *20*, 159-167.

van Elsas, J.D., Kastelein, P., van Bekkum, P., van der Wolf, J.M., de Vries, P.M., and van Overbeek, L.S. (2000). Survival of *Ralstonia solanacearum* Biovar 2, the Causative Agent of Potato Brown Rot, in Field and Microcosm Soils in Temperate Climates. Phytopathology *90*, 1358-1366.

van Noorden, G.E., Verbeek, R., Dinh, Q.D., Jin, J., Green, A., Ng, J.L., and Mathesius, U. (2016). Molecular Signals Controlling the Inhibition of Nodulation by Nitrate in *Medicago truncatula*. Int J Mol Sci *17*.

van Spronsen, P.C., Tak, T., Rood, A.M., van Brussel, A.A., Kijne, J.W., and Boot, K.J. (2003). Salicylic acid inhibits indeterminate-type nodulation but not determinate-type nodulation. Mol Plant Microbe Interact *16*, 83-91.

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D.*, et al.* (2018). Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proc Natl Acad Sci U S A *115*, E4700-E4709.

Vance, C.P. (2001). Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. Plant Physiol *127*, 390-397.

Vasse, J., Frey, P., and Trigalet, A. (1995). Microscopic Studies of Intercellular Infection and Protoxylem Invasion of Tomato roots by *Pseudomonas solanacearum*. Molecular Plant-Microbe Interactions *8*, 241-251.

Vavilin, D.V., and Vermaas, W.F. (2002). Regulation of the tetrapyrrole biosynthetic pathway leading to heme and chlorophyll in plants and cyanobacteria. Physiol Plant *115*, 9-24.

Vermeer, J.E., von Wangenheim, D., Barberon, M., Lee, Y., Stelzer, E.H., Maizel, A., and Geldner, N. (2014). A spatial accommodation by neighboring cells is required for organ initiation in Arabidopsis. Science *343*, 178-183.

Vernie, T., Kim, J., Frances, L., Ding, Y., Sun, J., Guan, D., Niebel, A., Gifford, M.L., de Carvalho-Niebel, F., and Oldroyd, G.E. (2015). The *NIN* Transcription Factor Coordinates Diverse Nodulation Programs in Different Tissues of the *Medicago truncatula* Root. Plant Cell *27*, 3410-3424.

Vernie, T., Moreau, S., de Billy, F., Plet, J., Combier, J.P., Rogers, C., Oldroyd, G., Frugier, F., Niebel, A., and Gamas, P. (2008). *EFD* Is an ERF transcription factor involved in the control of nodule number and differentiation in *Medicago truncatula*. Plant Cell *20*, 2696-2713.

Vidal, E.A., Alvarez, J.M., Moyano, T.C., and Gutierrez, R.A. (2015). Transcriptional networks in the nitrate response of *Arabidopsis thaliana*. Curr Opin Plant Biol *27*, 125-132.

Walker, L., Boddington, C., Jenkins, D., Wang, Y., Gronlund, J.T., Hulsmans, J., Kumar, S., Patel, D., Moore, J.D., Carter, A.*, et al.* (2017). Changes in Gene Expression in Space and Time Orchestrate Environmentally Mediated Shaping of Root Architecture. Plant Cell *29*, 2393-2412.

Wang, D., Yang, S., Tang, F., and Zhu, H. (2012). Symbiosis specificity in the legume: rhizobial mutualism. Cell Microbiol *14*, 334-342.

Wang, H., Ngwenyama, N., Liu, Y., Walker, J.C., and Zhang, S. (2007). Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in Arabidopsis. Plant Cell *19*, 63-73.

Wang, Q., Yang, S., Liu, J., Terecskei, K., Abraham, E., Gombar, A., Domonkos, A., Szucs, A., Kormoczi, P., Wang, T.*, et al.* (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. Proc Natl Acad Sci U S A *114*, 6854-6859.

Wang, Y., Penfold, C.A., Hodgson, D.A., Gifford, M.L., and Burroughs, N.J. (2014). Correcting for link loss in causal network inference caused by regulator interference. Bioinformatics *30*, 2779-2786.

Weidmann, S., Sanchez, L., Descombin, J., Chatagnier, O., Gianinazzi, S., and Gianinazzi-Pearson, V. (2004). Fungal elicitation of signal transduction-related plant genes precedes mycorrhiza establishment and requires the *DMI3* gene in *Medicago truncatula*. Mol Plant Microbe Interact *17*, 1385-1393.

Wildermuth, M.C., Dewdney, J., Wu, G., and Ausubel, F.M. (2001). Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature *414*, 562-565.

Withers, J., and Dong, X. (2017). Post-translational regulation of plant immunity. Curr Opin Plant Biol *38*, 124-132.

Wong, J., Nadzieja, M., Madsen, L.H., Bucherl, C.A., Dam, S., Sandal, N.N., Couto, D., Derbyshire, P., Uldum-Berentsen, M., Schroeder, S.*, et al.* (2019). A *Lotus japonicus* cytoplasmic kinase connects Nod factor perception by the *NFR5* LysM receptor to nodulation. Proc Natl Acad Sci U S A *116*, 14339-14348.

Xie, F., Murray, J.D., Kim, J., Heckmann, A.B., Edwards, A., Oldroyd, G.E., and Downie, J.A. (2012). Legume pectate lyase required for root infection by rhizobia. Proc Natl Acad Sci U S A *109*, 633-638.

Yamaguchi, Y., and Huffaker, A. (2011). Endogenous peptide elicitors in higher plants. Curr Opin Plant Biol *14*, 351-357.

Yamaguchi, Y., Huffaker, A., Bryan, A.C., Tax, F.E., and Ryan, C.A. (2010). *PEPR2* is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in Arabidopsis. Plant Cell *22*, 508-522.

Yamaguchi, Y., Pearce, G., and Ryan, C.A. (2006). The cell surface leucine-rich repeat receptor for *AtPep1*, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic Tobacco cells. Proc Natl Acad Sci U S A *103*, 10104-10109.

Yan, D., Easwaran, V., Chau, V., Okamoto, M., Ierullo, M., Kimura, M., Endo, A., Yano, R., Pasha, A., Gong, Y.*, et al.* (2016). *NIN-like protein 8* is a master regulator of nitrate-promoted seed germination in Arabidopsis. Nat Commun *7*, 13179.

Yanagisawa, S. (2014). Transcription factors involved in controlling the expression of nitrate reductase genes in higher plants. Plant Sci *229*, 167-171.

Yang, S., Tang, F., Gao, M., Krishnan, H.B., and Zhu, H. (2010). R gene-controlled host specificity in the legume-rhizobia symbiosis. Proc Natl Acad Sci U S A *107*, 18735-18740.

Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S.*, et al.* (2008). *CYCLOPS*, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci U S A *105*, 20540-20545.

Yao, J., and Allen, C. (2006). Chemotaxis is required for virulence and competitive fitness of the bacterial wilt pathogen *Ralstonia solanacearum*. J Bacteriol *188*, 3697-3708.

Yoro, E., Suzaki, T., Toyokura, K., Miyazawa, H., Fukaki, H., and Kawaguchi, M. (2014). A Positive Regulator of Nodule Organogenesis, *NODULE INCEPTION*, Acts as a Negative Regulator of Rhizobial Infection in *Lotus japonicus*. Plant Physiol *165*, 747-758.

Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F., Gouzy, J., Schoof, H.*, et al.* (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature *480*, 520-524.

Yu, L.H., Wu, J., Tang, H., Yuan, Y., Wang, S.M., Wang, Y.P., Zhu, Q.S., Li, S.G., and Xiang, C.B. (2016). Overexpression of Arabidopsis *NLP7* improves plant growth under both nitrogen-limiting and -sufficient conditions by enhancing nitrogen and carbon assimilation. Sci Rep *6*, 27795.

Zhang, C., Xie, Q., Anderson, R.G., Ng, G., Seitz, N.C., Peterson, T., McClung, C.R., McDowell, J.M., Kong, D., Kwak, J.M.*, et al.* (2013). Crosstalk between the circadian clock and innate immunity in Arabidopsis. PLoS Pathog *9*, e1003370.

Zhang, H., and Forde, B.G. (1998). An Arabidopsis MADS box gene that controls nutrient-induced changes in root architecture. Science *279*, 407-409.

Zhao, Y., Wang, T., Zhang, W., and Li, X. (2011). *SOS3* mediates lateral root development under low salt stress through regulation of auxin redistribution and maxima in Arabidopsis. New Phytol *189*, 1122-1134.

Zheng, Z., Qamar, S.A., Chen, Z., and Mengiste, T. (2006). Arabidopsis *WRKY33* transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J *48*, 592-605.

Zhu, Y., Dong, A., Meyer, D., Pichon, O., Renou, J.P., Cao, K., and Shen, W.H. (2006). Arabidopsis *NRP1* and *NRP2* encode histone chaperones and are required for maintaining postembryonic root growth. Plant Cell *18*, 2879-2892.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T., and Felix, G. (2006). Perception of the bacterial PAMP EF-Tu by the receptor *EFR* restricts Agrobacterium-mediated transformation. Cell *125*, 749-760.

Zolla, G., Heimer, Y.M., and Barak, S. (2010). Mild salinity stimulates a stress-induced morphogenic response in *Arabidopsis thaliana* roots. J Exp Bot *61*, 211-224.